

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 15:20:44 ; Search time 3927.32 Seconds
(without alignments)
5616.301 Million cell updates/sec

Title: US-09-488-265-25
Perfect score: 1426
Sequence: 1 tatatgaattcatggcggtg.....ttogcttaagaattcatata 1426

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_rod.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vl1.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rod1.*
95: gb_rod2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	% Match	Query Length	DB ID		
1	1342.8	94.2	1426	9	AX021809	AX021809 Sequence
2	790.8	55.5	1350	56	AF295325	AF295325 Synthetic
3	622.2	43.6	1404	9	A19452	A19452 phytase cDN
4	622.2	43.6	1404	10	I13430	I13430 Sequence cDN
5	622.2	43.6	1404	10	I33881	I33881 Sequence 19
6	619.6	43.5	1571	9	AX000634	AX000634 Sequence
7	619.6	43.5	1812	15	SFUS9804	U59804 Aspergillus
8	618.2	43.4	2000	13	ANPHYAG	Z16414 A.niger phy

9	618.2	43.4	2363	9	AR018076	Sequence
10	618.2	43.4	2363	9	AR051916	Sequence
11	618.2	43.4	2379	9	AR053934	Sequence
12	618.2	43.4	2379	13	ASNPHTAS	W02421 Aspergillus
13	618.2	43.4	2665	13	ASNPHTASE	M94550 Aspergillus
14	618.2	43.4	6756	9	AI9451	AY19451 phytase gen
15	618.2	43.4	6756	10	II13429	Sequence 31
16	613.4	43.0	1528	13	AF218813	AF218813 Aspergill
17	599.6	42.0	1515	12	AB022700	AB022700 Aspergill
18	597.6	41.9	1590	14	AY013315	AY013315 Aspergill
19	595.8	41.8	1553	12	AB042805	AB042805 Aspergill
20	580.4	40.7	1931	9	AX000630	Sequence
21	580.4	40.7	1931	14	ENU59803	U59803 Emericella
22	572.4	40.1	1912	9	A46793	Sequence 11
23	549	38.5	1845	9	AX000632	Sequence
24	549	38.5	1845	15	TU59802	TU59802 Talaromyces
25	533.4	37.4	1567	9	AX000636	AX000636 Sequence
26	533.4	37.4	1770	14	ATUG60412	U60412 Aspergillus
27	533.4	37.4	2327	9	A46783	AX46783 Sequence 1
28	533.4	37.4	2327	14	ATU59805	U59805 Aspergillus
29	512	35.9	1922	10	AX085191	AX085191 Sequence
30	510	35.8	1464	10	AX085193	AX085193 Sequence
31	510	35.8	1584	10	AX085192	AX085192 Sequence
32	412.4	28.9	4898	10	AX085207	AX085207 Sequence
33	355.6	25.0	2200	9	AR031151	AR031151 Sequence
34	292.6	20.5	853	10	AX085213	AX085213 Sequence
35	280.6	19.7	860	10	AX085208	AX085208 Sequence
36	265.2	18.6	3995	9	A46785	AX46785 Sequence 3
37	265.2	18.6	3995	14	TU59806	U59806 Myceliophth
38	134.4	9.4	950	10	MX085210	AX085210 Sequence
39	90.2	6.3	1117	14	MGR299239	AJ299239 Magnapor
40	69.6	4.9	1404	14	SCU19789	U19789 Saccharomyc
41	64.8	4.5	56097	15	YSC9177	U00029 Saccharomyc
42	64	4.5	1404	14	SCPHO5A	X01079 yeast gene
43	64	4.5	1904	10	AX072908	AX072908 Sequence
44	64	4.5	2934	15	SCYBR093C	Z35962 S.cerevisia
45	64	4.5	3913	14	SCPHO35	X01080 yeast genes

ALIGNMENTS

RESULT	1			PAT	07-SEP-2000
AX021809					
LOCUS	AX021809	1426 bp	DNA		
DEFINITION	Sequence 3 from Patent EP0897985.				
ACCESSION	AX021809				
VERSION	AX021809.1	GI:10045052			
KEYWORDS	. synthetic construct.				
SOURCE	synthetic construct				
ORGANISM	artificial sequence.				
	1 (bases 1 to 1426)				
REFERENCE	Lehmann, M.				
AUTHORS	Consensus phytaeses				
TITLE	Patent: EP 0897985-A 3 24-FEB-1999;				
JOURNAL	HOFFMANN LA ROCHE (CH)				
FEATURES	Location/Qualifiers				
source	1..1426				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="consensus sequence"				
	338 a 311 c 308 g 469 t				

```
Query Match          94.2%; Score 1342.8; DB 9; Length 1426;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
```

61	QY	catcgggtacgccttgggtcctcgtagtgaattctcaactcttgagcaactgttgacggtg	120
61	Db		
61	QY	CATCCGGTACCGCCTTGGGTCTCGTGGTAACTCTCACTCTTGTCACACTGTGTGACGGTG	120
61	Db		
121	QY	gttaccaatgtttccagaaaattctcaacttgggtggaataactctccattctctctt	180
121	Db		
121	QY	GTTACCAATGTTTCCAGAAATTTCTACATTGTGGGGTCAATACTCTCCATACTTCTCTT	180
121	Db		
181	QY	tggtgacgaactctgctattctccagcgttccaaaagggttgtagagttactttogttc	240
181	Db		
181	QY	TGGAAGACGAATCTGCTATTCTCCAGACGTTCCAGACGCTGTAGAGTTACTTTTCGTT	240
181	Db		
241	QY	aaqttttgtctagacacggtgctagataccacaactctctcaagtcctaagaagtactctg	300
241	Db		
241	QY	AAGTTTGTCTAGACACGGTGTAGATACCCAACTTCTCTAACTCTTAAGCTCTACTCTG	300
241	Db		
301	QY	ctttgattgaagctatcgaagaacgctactgctttcaagggttaagtcagctttcttga	360
301	Db		
301	QY	CTTTGATTGAAGCTATTTCAAAAGACGCTACTGCTTTCAAGGTAAGTACGCTTTCTTGA	360
301	Db		
361	QY	agacttaacaactactttgggtgctgacgacttgactccatctcggtgacacaagaatgg	420
361	Db		
361	QY	AGCTTACAACACTACTTTGGGTGCTGACGACTTGACTCTCATTCGGTGAACAACEAAATGG	420
361	Db		
421	QY	ttaactctggtattaaagttctacagaagatacaaggcttggctgagaagattgtccat	480
421	Db		
421	QY	TTAACTCTGGTATTAACTTCTACAGAAGATACAAAGCGTTTGGCTAGAAAGATTGTTCCAT	480
421	Db		
481	QY	tcggttagactctgggtctgtgacagattattgctctctctgctgaagaagttcatg	540
481	Db		
481	QY	TCATTAGAGCTTCCTGGTCTGACAGAGTTATTGCTTCTGTGAAAAAGTTCAATTGAAGGTT	540
481	Db		
541	QY	tccaactctgaagtgtggctgacccaggtctaacccacaacgctctccagttatta	600
541	Db		
541	QY	TCCAATCTGCTAAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG	600
541	Db		
601	QY	acgttattattccagaaggtgctggttacacaacaactttggaccacoggtttggtactg	660
601	Db		
601	QY	ACGTTATTATTCCAGAAGATCCGGTTACAACAACACTTTTGGACACGGTACTTGTACTG	660
601	Db		
661	QY	ctttcgaagaactgaaattgggtgacgagtgtaagtaacttcaactcgctgtttcgcctc	720
661	Db		
661	QY	CTTTGGAAGACTCTGAATTGGGTGACGAGTTGAAGCTTAATTCACACTTGTCTTTCGCTC	720
661	Db		
721	QY	caactattagactagattggaagctcacttgcaggtgttaaacttgactgacgaagaag	780
721	Db		
721	QY	CAGCTATTAGCGTAGATTGGAAGCTGACTTGGCCAGGTGTTACTTTGACTGACGAAGAC	780
721	Db		
781	QY	tgttaacttgatgacatgtgccaattgacactgttgtagaacttctgacgctactc	840
781	Db		
781	QY	TTGTTTACTTGATGGACATGTGTCCAATTCGAACACTGTGTGTAGAACTTCTCAGCGTACTG	840
781	Db		
841	QY	aattgtctcaattctgactgttcaactcacgacgaatggattcaatcacgactacttgc	900
841	Db		
841	QY	AATTGTCTCATTTCTGTGCTTTGTTCACTCACGACGAATGGAGACAATACGACTACTTGC	900
841	Db		
901	QY	aattcttgggtaagactacoggttacggtgctggtaaccattgggtccagctcaaggtg	960
901	Db		
901	QY	AATCTTTGGTAAAGTACTAGGGTTAGGGTGTGGTAAACCATTTGGTCCAGCTCAAGGTG	960
901	Db		
961	QY	tgggttcggttaacgaattgctagattgactcaactctccagttcaagaccacactt	1020
961	Db		
961	QY	TTGGTTTCGGCTAACGAATTGAATGCTAGATTGACTAGACTCTCCAGTTCAAGACCACACTT	1020
961	Db		
1021	QY	ctactaaccacactttggaactctaaaccagctactttccattgaacgctactttgacg	1080
1021	Db		
1021	QY	CTACTAACACACTTTGGACTTAACCCAGCTACTTTCCTCATTTGACAGCTACTTGTACG	1080
1021	Db		
1081	QY	ctgactctctcagacaacaactatggttctattttcttcgcttgggtttgtacacag	1140
1081	Db		
1081	QY	CTGACTCTCTCAGACAACTCTATGATTTCTATTTTTCTTCGTTTGGTTCGTACACG	1140
1081	Db		

[illegible]

SPFCOLTHNEWKYNYLSQIKYGYGAGNPLGPAOGIGFTNELIARLTPSPVDHT
 STNLSYNSPATPLNATMYVDESHNSMVSIFFAIGLYNGTEPLSPRTSVESAKELDG
 YSASVVVFGARAYFETMOCKSEKEPLVRLINDRVPLHGDVDLGRCKLNDFFVKG
 LSWARSGNNGEFS"
 BASE COUNT 326 a 416 c 453 g 376 t
 ORIGIN

Query Match 43.5%; Score 619.6; DB 9; Length 1571;
 Best Local Similarity 67.0%; Pred. No. 7.4e-154;
 Matches 896; Conservative 0; Mismatches 439; Indels 3; Gaps 1;
 QY 86 tggtaattctcaactcttggtaactgtgaacgtgtgttaaccatgtttccagagaaatttc 145
 DB 171 TGTGCTGCTCAAGTCTGCGATACGGTAGACCTCGGGGTACCAAGTGTCCCTCGGACTTC 230
 QY 146 tcaacttggtgggttaact 205
 DB 231 TCATCTATGGGGCCAGTACTTCGCCATCTCTTTCGCTCGAGGACGACTCTCCGTGTCCAG 290
 QY 206 agacgttccaaaggggttagagttacttctctctctctctctctctctctctctctctctct 265
 DB 291 TAAGCTTCCCAAGATTGCGGATACCTTGTACAGCTGTATCGCGCATGGAGCGC 350
 QY 266 ataccacattcttctaagttctaagtaactctctctctctctctctctctctctctctctct 325
 DB 351 GTACCCCAACAGCTCCAGAGCAAAAGTATAAGAAAGCTGTGACGGCGATCCAGGCCAA 410
 QY 326 cgtctactgtttcaaggtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 385
 DB 411 TGCACCGGACTTCAAGGCAAGTTGGCTTTTGAAGACGCTACAACCTATCTCTGGGTGC 470
 QY 386 tgaactgactctcattcgttgaacacaaatggttaactctgttatttaagttctacag 445
 DB 471 GGATGACCTCACTCCTTTGGGAGCAGCAGCTGTGNACTCGGGCATCAAGTTCTACCA 530
 QY 446 aagatacaagcttggtagaagattgttccattgttagagcttctgttctgtctgtacag 505
 DB 531 GAGGTACAGGCTCTGGCGCCAGCTGGTGGCTTTATTTCGCGCTCAGGCTCGGACCG 590
 QY 506 agttattgtctctgtctgaagttcattgaagtttccaatctgtctgaagttgtgacct 565
 DB 591 GGTATTGCTTCGGAGAGAGTTCATCGAGGGTTCACAGCGGAGTCCAGCGGAGCTGGCTATCC 650
 QY 566 aggtcttaaccac 625
 DB 651 TGGCGCGA---CGAACCGCGCGCTCGCGGATAGTGTGATATTCCGAGAGCAGAC 707
 QY 626 ttacaac 685
 DB 708 GTTCAACAATACGCTGGACCCAGCGTGTGTCACGAAGTTTGAGCGCAGTCAGCTGGGAGA 767
 QY 686 cgacttgaagcttaactctactgtcttctcgtctccacacattagagctagattggaagc 745
 DB 768 TGAGGTTCGGGCAATTTCACTCGCTCTTTGACCCGACATCCGAGCTCGCGCCGAGAA 827
 QY 746 tcaactggcagggttaactgaacgaagacgtgttgaactgtgacatgtgtcc 805
 DB 828 GCATCTCTCTGCGCTGACCTGACAGACGAGGACGCTGTGCTCAATGGACATGTGTC 887
 QY 806 attcgactgttctgaactctctgaactactcaactcaacttcttgggaagtaactacggcta 865
 DB 888 GTTTGATAGGTAAGCGGACCGACGACGACGACGACGACGACGACGACGACGACGACGAC 947
 QY 866 caactcacgaatgagttcaactactactactactactactactactactactactactactact 925
 DB 948 CACTCACATGAGTGAAGAAGTACAACCTCACTCACTCACTCACTCACTCACTCACTCACT 1007
 QY 926 cgggtgtgttaaccattgggtccagctcaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 985
 DB 1008 CGGCGCAGCAACCTCTGGGACCGGCTCAGGGGATAGGGTTACCAACGACGCTGATTGC 1067
 QY 986 tagattgactcaactctcagttcaagaccacactctactaaccacacacttggacttaa 1045

DB 1068 CCGGTTGACTCGTCCGAGTCCAGGACACACACGACTTAACCTGACTCTAGTCTCCAA 1127
 QY 1046 ccagctacttccactgaacgtacttctgtacgctgacttctctcagcagcaacactat 1105
 DB 1128 CCGGCGCACTTCCCGTTGAACGCTACCATGTACGTCGACTTTTCACAGCAACAGCAT 1187
 QY 1106 ggtttctatttctcgtcttgggtttgttatacagcgttactaagccattgttactacttc 1165
 DB 1188 GGTTCCTCATCTCTTGGATGGGCTGTACAAAGGCACTGAACCCCTTGTCCCGGACCTC 1247
 QY 1166 tttgaactattgaagaaactgacggttgcgtctctctctctctctctctctctctctct 1225
 DB 1248 GGTGAAAGCGCCAGGAATGGATGGGTATTCGATCTCTGGTGGTGGTCTTTCGGCGC 1307
 QY 1226 tagagcttcaactgaatgacggttgcgtctctctctctctctctctctctctctctctct 1285
 DB 1308 CGGAGCTTACTTCAGAGCATGCAATGCAAGTCGAAAGGAGCCTCTGTTCGCGCTTT 1367
 QY 1286 ggttaacgacagagttcttccattgcacggttgcgtctctctctctctctctctctctct 1345
 DB 1368 GATTAAATGACCGGTTGTGCACATGCTGCTCGATGTCGATGTCGACAAAGCTGGGCGATGCAA 1427
 QY 1346 gagagacgactctgtgaagtttctctctctctctctctctctctctctctctctctct 1405
 DB 1428 GCTCAATGACTTTGTCAAGGATGTAGTTGGGCGCAGATCTGGGGGCAACGAGGAGATG 1487
 QY 1406 tttcgtttaagaattcat 1423
 DB 1488 CTTTAGTTGAGATGTCAT 1505

RESULT 7
 SFU59804
 LOCUS SFU59804 1812 bp DNA PLN 17-JAN-1998
 DEFINITION Aspergillus fumigatus phytase gene, complete cds.
 ACCESSION U59804
 VERSION U59804.1 GI:2108353
 KEYWORDS
 SOURCE Aspergillus fumigatus.
 ORGANISM Aspergillus fumigatus
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE 1 (bases 1 to 1800)
 AUTHORS Pasamontes,L., Haiker,M., Wyss,M., Tessier,M. and van Loon,A.P.
 TITLE Gene cloning, purification, and characterization of a heat-stable
 phytase from the fungus Aspergillus fumigatus
 JOURNAL Appl. Environ. Microbiol. 63 (5), 1696-1700 (1997)
 MEDLINE 97288063
 REFERENCE 2 (bases 1 to 1812)
 AUTHORS Pasamontes,L.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1996) Luis Pasamontes, VFCEB, F. Hoffmann-La Roche
 AG., Basel 4070, Switzerland
 FEATURES
 Location/Qualifiers
 1..1812
 /organism="Aspergillus fumigatus"
 /db_xref="taxon:5085"
 195..198
 TATA_signal
 197..200
 TATA_signal
 join(284..330,388..418)
 sig_peptide
 join(<284..330,388..>1738)
 join(284..330,388..1738)
 CDS
 /EC_number="3.1.3.8"
 /codon_start=1
 /product="phytase"
 /protein_id="AAB96872.1"
 /db_xref="GI:2108354"

/translation="MVTITFLLSAAVLLSGRVSAAPSSAGSKSCDVTDLGTQCSPTS
 HLWGQYSPFSLDELSSVKLPKDRITLVQLVLSRHGRIPTSSKSKKKYKLLTAIQ
 ANATDKGKFAFLKTYNTIGADDLTPFGEOQLVNSGKFKYORYKALARSVYVFTIRAS
 GSDYFIASGEKFIETGFOAKLADPGATNRAAPAIISVIIPESETFNNTLDHGVCKFEA
 SOLGDEVAANTFALFAPDIRAEKHLPCVTLTDEDVVVSLMDMCGSFDTVARTSDASQL

Sat Oct 27 15:25:36 2001

SPFCQFTHNEWKYNILQSLGKYYGYGAGNPLGPAQIGTINELIARLSPVDHT
 YSNLVSNPATFPLNATMYDYFSDHNSMWSIFFALGLYNGFLEPLSRTSVESAKELDG
 YSASWVPFGARAYFETWOCKSEKPELVRALINDRVPLVPLHGCDVDKLGKRLKNDLVKG
 LSWARSGNNGWCEFS"
 BASE COUNT 377 a 480 c 532 g 423 t
 ORIGIN

Query Match 43.5%; Score 619.6; DB 15; Length 1812;
 Best Local Similarity 67.0%; Pred. No. 7.5e-154;
 Matches 896; Conservative 0; Mismatches 439; Indels 3; Gaps 1;

Qy 86 tggtaattctcaactcttctgtgacactgttgacggtgttaccactgtttcccaagaatttc 145
 Db 412 TGTGGCTCCAAAGTCCCGATACGGTAGACCTCGGGTAGCAGTGTCCCTCGGACTTC 471
 Qy 146 teacttgggggtcaataactctccactctctctcttctgtgacgaatctgctattctcc 205
 Db 472 TCATCTATGGGGCCAGTACTCGCCATCTCTTCGTCGAGAGAGAGTGTCCGTGTCGAG 531
 Qy 206 agacgttccaaaggtgttagagttacttcttcttcttcttcttcttcttcttcttct 265
 Db 532 TAACTCTCCCAAGGATTCGGCGATACCTTGGTACAGGTCTATFCGGCGCATGGAGCGG 591
 Qy 266 ataccactcttcttaagtctaagaagtactctgttcttcttcttcttcttcttcttct 325
 Db 592 GTACCCACACGCTCCAGAGCAAAAAGTATTAAGAAGCTTGTGACGCGCATCCAGGCCAA 651
 Qy 326 cgtactgttttcaaggtgaagtacgtcttcttcttcttcttcttcttcttcttcttct 385
 Db 652 TGCACCGACTTCAAGGCAAGCTTTGGCTTTTAAAGACGTTACAACTATATCTCFGGGTG 711
 Qy 386 tgacgtactgtactcaattctgttgacacaaagtgttaactctgtgttattgaattctacg 445
 Db 712 GGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGTGAACCTCGGCATCAAGTCTACCA 771
 Qy 446 aagatacaaggtcttggcttagaagaattgttccattcttcttcttcttcttcttcttct 505
 Db 772 GAGTACAGGCTCTGCGCGAGTGTGTGGCTTTATTCGCGCTCAGGCTCGGACCG 831
 Qy 506 agttattgtctgtgaaagtcttgaagtttcaactcttcttcttcttcttcttcttcttct 565
 Db 832 GGTATTGCTTCGGGAGAGAAGTTCATCAGGGGTTCAGACGGGAGAGCTGGCTGATCC 891
 Qy 566 agtgcctaaccacacacaaagcttctccagttattacttcttcttcttcttcttcttctt 625
 Db 892 TGGCGCA---CGAACCGCGCTCCGCGATTAGTGTGATTATTCGCGAGAGCGAGAC 948
 Qy 626 ttacaacaaactttgacacacggttctgtactgtcttcttcttcttcttcttcttcttct 685
 Db 949 GTTCAACAATACGCTGACACCGGTGTGTGCACGAAGTTTGAGGCGAGTCAGCTGGAGA 1008
 Qy 686 cgagctgaagcttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 745
 Db 1009 TGAGTTGCGGCCATTTCACTGCGCTTTTGGACCGCATATCCGAGCTCGCGCGAGAA 1068
 Qy 746 tcaacttgcaggtgttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 805
 Db 1069 GCATCTTCTGCGGTGAGCTGTGACAGACAGGAGCTTGTCTCAGTCTATGACATGTGTC 1128
 Qy 806 attcagactgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 865
 Db 1129 GTTTGATAGGTAGGCGGCACAGCGCAAGTCAGCTGTCACTGCTGTCTCAACTCTTT 1188
 Qy 866 cactcagacgaattgattcaactacttcttcttcttcttcttcttcttcttcttcttcttctt 925
 Db 1189 CACTCACAATGAGTGAAGAAGTACAACTCTTCACTCTTCTGCGCAAGTACTACGSCTA 1248
 Qy 926 cgttactgttaaccattgggttccagcttcaaggtgttgggttcttcttcttcttcttctt 985
 Db 1249 CGCGCAGCAGCACTCTTGGGACCGCTCAGGGGATAGGTTTCAACCAACGAGCTGATTC 1308
 Qy 986 tagattgactcactctccagtttcaagacacacacttcttactaaccacttttggactctaa 1045

RESULT 8

ANPHYAG ANPHYAG 2000 bp DNA PLN 06-OCT-1992
 LOCUS A.niger phyA gene.
 DEFINITION Z16414
 ACCESSION Z16414.1 GI:2392
 VERSION phyA gene.
 KEYWORDS Aspergillus niger.
 SOURCE Aspergillus niger.
 ORGANISM Aspergillus niger.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE 1 (bases 1 to 2000)
 AUTHORS van Hartingsveldt, W., Van Zeijl, C. M. J., Harteveld, M. G., Gouka, R. J.,
 Suykerbuyk, M. B. G., Luiten, R. G. M., Van Paridon, P. A., Selden, G. C. M.,
 Veenstra, A. E., Van Gorcom, R. F. M. and Van Den Hondel, C. A. M. J.
 Cloning, molecular characterization and overexpression of the
 phytase gene (phyA) of Aspergillus niger
 JOURNAL Gene (1992) In press
 REFERENCE 2 (bases 1 to 2000)
 AUTHORS van Hartingsveldt, W.
 TITLE Direct Submission
 JOURNAL Submitted (05-OCT-1992) Van Hartingsveldt W., TNO Medical
 Biological Laboratory, Lange Kleiweg 139, Rijswijk, the Netherlands
 FEATURES
 source
 1. .2000
 /organism="Aspergillus niger"
 /strain="Van Tiegheem strain"
 /db_xref="taxon:5061"
 TATA_signal 131. .134
 exon 166. .253
 /gene="phyA"
 /number=1
 166. .1871
 /gene="phyA"
 /partial
 prim_transcript 166. .1871
 /genes="phyA"
 /function="primary transcript of phyA gene"
 join(210. .253, 356. .365)
 sig_peptide /gene="phyA"
 CDS join(210. .253, 356. .1715)

```

/gene="phyA"
/codon_start=1
/protein_id="CAA78904.1"
/db_xref="GI:2393"
/translators="MGVSAVLLPLLYLLSGVTSGLAVPASRNSSDDVDVQGYOCFSET
SHLWGYAPFPFSLANESVSPVACGRVTPAQVLSRHGARYPTDSKGYKYSALIEEI
QONATTFDKGYAPLKTNYSLGADDLTPFGEQELVNSGKIFORYESLRTNTPVPIRS
SGSSRVASCKKFEIEGFSQTKLKDPAQPGSQSPKIDVVYISEASSNNTLDPTCTVF
EDSELADTEANFTVFPVSIORLENDLSGVTLDTTEVYLMDCMSFDTLSTSDVT
KLSFCDLTHEDWINYDLSKLYYGHGAGNPLGPTQGVYANELLARLHSPVHD
DYSNHTLSDSPATFPNLSYVADESHDNGIISILFALGLYNGTPLSLTIVENITOT
DGFSSAWTVPFASRLYVENMQCAQEPFLVRVLVNDVRVPLHGCPVDALGRCRDRSFV
RGLSFARSGCDWAECFA"
254..355
intron
/gene="phyA"
/number=1
356..1878
exon
/number=2
369..1712
mat_peptide
1732..1760
stem_loop
/gene="phyA"
BASE COUNT 463 a 603 c 456 g 478 t
ORIGIN

```

```

Query Match 43.4%; Score 618.2; DB 13; Length 2000;
Best Local Similarity 65.3%; Pred. No. 1.8e-153;
Matches 908; Conservative 0; Mismatches 483; Indels 0; Gaps 0;
QY 24 gtctgtactgtccatgcccaccttggctcggtccacatccggtacgcgcttgggtctt 83
DB 324 GTGGGACTACTGATCGCTGACTATCTGTGCAGAGTCACCTCCGAGTGGCAGTCCCGGCC 383
QY 84 cgtgttaattcactctgtgacactgtgtgacggtgtgtaccgaatgtttcccaagaatt 143
DB 384 TCGAGAAATCAATACAGTTCGATACGCTGCATCAGGGGTATCAATGCTCTCCGAGACT 443
QY 144 tctcaacttgggttcaactactcctccactcttcttcttggctgacgaatctgctattct 203
DB 444 TCGCATCTTTGGGGTCAATACCGACCGCTTCTCTCTGTCGCAAAACGAATCGGTCACTCC 503
QY 204 ccagacgttccaaaggtgttagagttacttctcggttcaagtgttctgctagacacggtgct 263
DB 504 CCTGAGGTGGCCGGGATCCAGAGTCACATTCGCTCAGGTCTCTCCGCTCATGAGCG 563
QY 264 agatacccaacttcttaagttaagaagtactctgtctgttgaagtacttcaaaag 323
DB 564 CGGTATCCGACCGACTCCAAAGGCAAGAAATACTCCGCTCTCATTGAGGAGATCCAGCAG 623
QY 324 aacgtactgtcttcaaggttaagtaacgcttcttgaagacttacaactacactttgggt 383
DB 624 AACCGGACCACTTTGACGGAAATATGCGCTTCTCTGTCGCAAAACGAATCGGTCTGGCT 683
QY 384 gctgacgacttactcctcgtgtaacacaaatgttgaactctggtattgaagtctac 443
DB 684 GCAGATGACGTACTCCCTTCGAGAACAGAGGAGTAGTCACTCCGCGCATCAAGTCTTAC 743
QY 444 agaagatacaggttggctgagaagattgttccattcgttagagcttctggttctgac 503
DB 744 CACCGGTAGATCGCTCAAGAGGAACATCGTTCATTATCCGATCTCTGGCTCCAGC 803
QY 504 agagttattgtcttctgtaaaagttcattgaaggtttccaatctgtaagtgtgctgac 563
DB 804 CGCGTGATCGCTCCGGCAAGAAATTCATCGAGGCTTCCAGAGCACCACCAAGTGAAGAT 863
QY 564 ccaggtgtaaccacacaaagcttctccagttattaaactgtattattccagaaggtgct 623
DB 864 CTCGTGCGCCAGCCGGCCAAATCGTCGCCCAAGATCGACGTGGTCAATTCGAGGCCAGC 923
QY 624 ggttacaacaacactttggaccacggttgggtactgcttctcgaagaatctgaattgggt 683
DB 924 TCAATCCAAACACACTCTCGACCCAGGCACCTGCACCTGTCTTCTGAAGACAGCAATTGGCC 983
QY 684 gacgagcttgaagcttaacttcaactcgtgttttcgctccactattagagctagattgaa 743
DB 984 GATACCGTTCGAGCCAAATTTACCCGCCAGGTTCTGCCCTCCATTCGTCACAGCTCTGGAG 1043
QY 744 gctcacttgcaggtgtttaaacttgaactgacgaagacgttgttaacttgaatgacatgct 803
DB 1044 AACGACCTGTCCGGTGTGACTCTCAGACACACAGAGTGACCTACCTCCTGACATGACATG 1103
QY 804 ccattcgcactgttctagaaacttctgaactcactcgaactcgaacttctccttctgacttg 863
DB 1104 TCCTTCGACACCATCTCCACGACGACCGCTCGACACCAAGCTGCCCTCTCTGTGACCTG 1163
QY 864 ttaactcagcgaatggattcaatcagactacttgaacttgaacttcttgggttaactactacggt 923
DB 1164 TTCACCCATGACGAATGATCACTACGACTACCTTCAGTCCCTTTGAAAAAGATTACGGC 1223
QY 924 taaggtgctgtaaccacttgggtccaggttccaggtcagcagcagcagcagcagcagcagc 983
DB 1224 CATGTCAGGTAACCGCTCGGCCGACCCAGCCAGCGCTCGGCTACGCTAACGAGCTCATC 1283
QY 984 gctagattgactcacttccacttcaagacacacacacacacacacacacacacacacacac 1043
DB 1284 GCCGCTGACGACACTCCGCTGTCACAGATGACACACAGTTCACACACACACTTTGGACTCG 1343
QY 1044 aaccagctacttccacttgaacgctacttctgacgctacttgaacgctacttctcagcacaac 1103
DB 1344 AGCCGGCTACTCTTCGGCTCACTCTACTCTCTACCGGACTTTTCGATGATCAACAGCGC 1403
QY 1104 atggttcttatttcttctgcttgggttgggttgggttgggttgggttgggttgggttgggt 1163
DB 1404 ATCATCTCCATCTCTTGTCTTAGTCTGTACACGGCAGTAAAGCCGCTATCTACACAG 1463
QY 1164 tctgttgaaacttattgaagaacacacacacacacacacacacacacacacacacacacac 1223
DB 1464 ACGGTGAGNAATATACCCAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1523
QY 1224 gtagagcttacttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1283
DB 1524 TCGCGTTTCTAGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1583
QY 1284 ttggttgaacacacacacacacacacacacacacacacacacacacacacacacacacac 1343
DB 1584 TTGGTTAATGATGCGGCTGTCCCGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1643
QY 1344 aagagacgacttctggttgaaggttcttctcgttagatctcgttagatctcgttagatctc 1403
DB 1644 ACCCGGATAGCTTTGTCAGGGGTTGAGCTTCTGCTAGATCTGGGGTGTGATTGGCGGAG 1703
QY 1404 ttttctcgtta 1414
DB 1704 TGTCTTGTCTTA 1714

```

```

RESULT 9
AR018076
LOCUS
DEFINITION Sequence 7 from patent US 5780292.
ACCESSION AR018076
VERSION AR018076.1 GI:3973679
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2363)
AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
Torikell,T.K., Cantrell,M., Piddington,C.S., Rambossek,J.A.,
Turunen,M.K. and Fagerstrom,R.B.
TITLE Production of phytate degrading enzymes in trichoderma
JOURNAL Patent: US 5780292-A 7 14-JUL-1998;
FEATURES Location/Qualifiers
source 1..2363

```

[illegible]

Db 698 CTGATGTCGCCGCGGTTGACAGAGTCACTTTCGCTCAGGTCTCTCCGCTCAATGAGCG 757
 Qy 264 agatacccaactcttcttaagtctaaagaactactctgtttgattgaagctattcaaaag 323
 Db 758 CGGTATCCGACCGAGTCAAGGGCAAGAAATACCTCCGCTCTCTATGAGGATCCAGCAG 817
 Qy 324 aacgactactgttccaaaggttaagtaagcttcttgaagacttacaactacacattgggt 383
 Db 818 AACGTGACCACTTTGATGGAATAATGCTTCTCTGAAGACATACACTACAGCTTGGGT 877
 Qy 384 gctgacgactgactccattcgtgaacacaaatgttaactctgttattaaagtctac 443
 Db 878 GCAGATGACCTGACTCCCTTCGGAGAGCAGAGAGTCACTAAGTCAACTCCGCTCAAGTTCTAC 937
 Qy 444 agaagatacagaagcttgggtgtagaagattgttccattcgttgaagactctgtttctac 503
 Db 938 CAGCGATACGAATCGCTCAACAGGAACATCATCTCCGTTCAATCCGATCCCTCTGGCTCCAGC 997
 Qy 504 agagttatgtctctgctgaaggttcatgaaggtttccaatctgctaagtgtgctgac 563
 Db 998 CGGTGATCGCTCCGCGGAGAAATTCATTTAGGGCTTCAGAGCACCAGCTGAAGAT 1057
 Qy 564 caaggtgttaacccacacacacactcttcagttattaaagttattattccagaaggtgct 623
 Db 1058 CCTGTGCGCCAGCGGGCAATCTGCGCCAGATCGAGGTGCTTCCGAGGCCAGC 1117
 Qy 624 ggttaacacacacttggagacacggttggtagctgttctgaagaactctgaattggt 683
 Db 1118 TCATCCAAACACACTCTCGACCCAGGACCTGCACCTGCTTTGAAGACAGCAATGGCC 1177
 Qy 684 gacgacttgaagctaaactcactgctgttttcgctcactattagagctagattgaa 743
 Db 1178 GATACCGTCGAAGCAATTTACCGCCACGTTTCGCCGCCCTCCATTCGTCAACGCTGGAG 1237
 Qy 744 gctcaactgcaaggtgttaactgactgacgaagacgtgttgaacttgatgacattggt 803
 Db 1238 AACGACCTGTGTGGCGTACTCTCACAGACACAGAAGTGACCTACCTCAATGGACATGTC 1297
 Qy 804 ccattgacactgtgttagaactctgacgactactcaattgtctcattctgtgacttg 863
 Db 1298 TCCCTGCACACCATCTCCACGACGCGTGCACACCAAGGTGCTCCCTCTGTGACCTG 1357
 Qy 864 ttaactcagcagcaagtgatttcaatacagactacttgaacttgaacttgaacttgaact 923
 Db 1358 TTCACCCATGACGAATGGATCCACTACGACTACCTCCAGTCCCTGAAAATACTACGGC 1417
 Qy 924 taagggtgtgaacccattgggtccagctcaaggttgggttgggttgaacttgaacttgaact 983
 Db 1418 CATGGCGCAGGTAAACCGCTCGGCGCGACCCAGGGGTGCGGTACGCTAACGAGCTATC 1477
 Qy 984 gctagattgactcactctccagttcaagaccacacttctactaaccacactttgactct 1043
 Db 1478 GCCCGTCTACCCACTCGCTGTCCAGATGACACAGCTCCACACACACTTGGACTCG 1537
 Qy 1044 aaccagctacttccattgaacgacttcttgaacttgaacttgaacttgaacttgaact 1103
 Db 1538 AACCGAGCTACCTTCCCGTCAACTCTACTCTCTACGCGGACTTTCCACACATACGCGC 1597
 Qy 1104 atggtttctatttcttgccttgggttgggttgaacttgaacttgaacttgaacttgaact 1163
 Db 1598 ATCATCTCTATCTCTTGTGGTCTGTACAAAGGACCTTAAGCCGCTGTCTACACAG 1657
 Qy 1164 tctgttgaactctattgaagaaactgacggttgaacttgaacttgaacttgaacttgaact 1223
 Db 1658 ACCGTGGAGAAATATACCCAGACAGATGGGTCTCGTCTGTGGACGTTCCGTTTGGT 1717
 Qy 1224 gctagagcttaacttgaatgatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1283
 Db 1718 TCGCTCTGTAGTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
 Qy 1284 ttggttaacgacagagttgttccattgacggttgggttgggttgggttgggttgggttgggt 1343

Db 1778 TTGGTTAATGATCCGCTTGTCCGCTGCAAGGGTGTCCAATTCATCTTTGGGAGATGT 1837
 Qy 1344 aagagagacgacttgggtgaaggttcttctcgtatgacttgggttaacttgggttaact 1403
 Db 1838 ACCCGGATAGCTTGTGTGAGGGGTGTAGCTTTGCTAGATCTGGGGTCAATTTGGCGGAG 1897
 Qy 1404 tgtttcgctta 1414
 Db 1898 TCTTCTGCTTA 1908
 RESULT 11
 AR053934
 LOCUS AR053934 2379 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5834286.
 ACCESSION AR053934
 VERSION AR053934.1 GI:5978796
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2379)
 AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Fagerstrom,R.B.,
 Miettinen-Oinonen,A.S.K., Turunen,M.K., Rambosek,J.A.,
 Piddington,C.S., Houston,C.S. and Cantrell,M.A.
 TITLE Recombinant cells that express phytate degrading enzymes in desired
 ratios
 JOURNAL Patent: US 5834286-A 1 10-NOV-1998;
 FEATURES Location/Qualifiers
 SOURCE l. .2379
 BASE COUNT 561 a 735 c 518 g 565 t
 ORIGIN
 Query Match 43.4%; Score 618.2; DB 9; Length 2379;
 Best Local Similarity 65.3%; Pred. No. 1.8e-153;
 Matches 908; Conservative 0; Mismatches 483; Indels 0; Gaps 0;
 Qy 24 gtcgtactgtccattgcccacccctgttcggttccacatccggtaccgcttgggtct 83
 Db 534 GTGGGACTACTGATCGTGACAACTGTGTCAGAGTACCTCCGGACTGGCAGTCCCGCC 593
 Qy 84 cgtgtaattctcactctgtgacactgttgacggtgttaccagtgttcccaatttcccagaatt 143
 Db 594 TCGAATAATCAATTCACCTTGCATACGCTGATACGATCAAGGTATCAATGCTTCTCCGAGACT 653
 Qy 144 tctcactgtgggttcaatactctcactctccttctcttcttggctgacgaactctgtctattct 203
 Db 654 TCGCATCTTTGGGGTCAATACGCGCCGCTTCTCTCTGTGGCAACGAATCGGCCATCTCC 713
 Qy 204 ccagactgtccaaaggtgttagagttacttctcctcaagtgttcttagacacggtgct 263
 Db 714 CCTGATGTGCGCGCGGTTGACAGAGTCACTTCTGCTCAGGTCCTCTCCGCTCATGGAGCG 773
 Qy 264 agatacccaacttcttcaagtctaaagtactctgcttggattgaagctatttcaaaag 323
 Db 774 CGGTATCCGACCGAGTCCAGGGCAAGAAATACCTCCGCTCTCATTTGAGGAGATCCAGCAG 833
 Qy 324 aacgctactgttccaaaggttaagtagcgttcttgaagacttacaactacactttgggt 383
 Db 834 AACGTGACCACTTTGATGGAATAATGCTTCTCTGAAGACATACACTACAGCTTGGGT 893
 Qy 384 gctgacgacttgcctcattcgtgaacacaaatgttccattcgttgaagcttctgatttaagttctac 443
 Db 894 GCAGATGACCTGACTCCCTTCGGAGAGCAGAGCTAGTCAACTCCGCGCATCAAGTTCTAC 953
 Qy 444 agaagatacagaagcttgggtgtagaagattgttccattcgttgaagcttctgatttctgac 503
 Db 954 CAGCGATACGAATCGCTCAACAGGAACATCATCTCCGTTCAATCCGATCTCTGGCTCCAGC 1013
 Qy 504 agagttatgtctctgctgaaggttcaatgaggttctccaatctgctaagtgtgctgac 563

		Eurotiales; Trichocomaceae.mitosporic Trichocomaceae; Aspergillus.
	REFERENCE	1 (sites)
	AUTHORS	Piddington,C.S., Houston,C.S., Paloheimo,M., Cantrell,M., Miettinen-Oinonen,A., Nevalainen,H. and Rambosk,J.
	TITLE	The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var. awamori
JOURNAL		Gene 133 (1), 55-62 (1993)
MEDLINE		94040796
REFERENCE		2 (bases 1 to 2379)
AUTHORS		Carter,J.R., Franden,M.A., Aebersold,R.H. and McHenry,C.S.
TITLE		Molecular cloning, sequencing and overexpression of the gene encoding the psi subunit of E. coli DNA polymerase III holoenzyme Unpublished (1992)
JOURNAL		Location/Qualifiers
FEATURES	source	1..2379 /organism="Aspergillus niger" /db_xref="taxon:5061"
	exon	420..463 /citation=[2] /evidence=experimental /join(420..463,566..1925) /ec_number="3.1.3.8"
	CDS	/citation=[2] /codon_start=1 /function="conversion of phytate to inositol and inorganic phosphorus" /evidence=experimental /product="phytase" /protein_id="AAAI16898.1" /db_xref="GI:166519"
		/translation="MGVSALLPLIYLVAGVTSLGVLPASRNOSTCDTVQGYCFSET SHLWGOAPFESLANESAIPDPVACRVTFQVLGRHARPYTESKHKYSALTEEI SQNVTPDGKYAFLEKTYNSLGADDLTPFGQEELVNSGKFYORYESLRNLIPIERS QGSNRVASGEKFTGFQSTKLKPRAQDQSSPKIDVISEASSNNILDDPTCIV EDSELADPEANETATPAPSIRORLENDLSGVTITDTETVLYMDMCSFDTISTVDY KLUSPDLFTHDEWHYDYLOSLAKYYHGAGNPGLPTQGVGVANELLIALRTHSPVHD DTSSNHLLDNSNPATFPNLSTLYADFSDHNGIIISILFALYNGTKPLSTTTVENITQT DGFSAANTVPFAASRLYVEMMOCAEQEPLVRVLVDRVPLHGCPTDALGRCRTDRSFV RGLSFARSGGDWAECSA"
	exon	566..1925 /note="putative" /citation=[2]
BASE COUNT	561 a 735 c 518 g 565 t	
ORIGIN		
	Query Match	43.4%; Score 618.2; DB 13; Length 2379;
	Best Local Similarity	65.3%; Pred. No. 1.8e-153;
	Matches	908; Conservative 0; Mismatches 483; Indels 0; Gaps 0;
QY	24	glcgtgtactgtccattgcacactgttcggttcacacatccgggtaccaccttgaggtoct 83
Db	534	GTCGGACTTACTGTATCGCTTGACAATCTGTGCAGAGTCACCTCCGGACTGCCAGTCCCGGC 593
QY	84	cgtggaatttcactctctgtgacactgttgtgacagtggttgatcaaatgtttccagaagt 143
Db	594	TCGAGAATAATCAATCCACTTCGGCATACGTCATCAAGGGTATCAATGCTTCTCCGAGACT 653
QY	144	tctcactgtgggttaataactctccattctctctgtgctgaacaaatgtttccagaagt 203
Db	654	TCGCATCTTTGGGGTCAATACGCCCGCTTCTCTCTCTGGCAAAGGAATCGGCATCTCC 713
QY	204	ccagacgttccaaaagggttagagttactttcgtaaagttttgttcacacacagtgct 263
Db	714	CCTGATGTGCCCGCGGTTCGAGAGTCATTTCGCTCAGTCTCTCCGCTCATPGAGCG 773
QY	264	agataccaactctcttaagtctaagaagtactctgtgttgatgaagctattcaaaag 323
Db	774	CGGTATCCGACCGAGTCCAAGGGCGGAAGAAATACTCCGCTCTCATTTAGGAGATCCAGCAG 833
QY	324	aacgctactgctttccaagggtgaagtacgctttcttgaagacttacacactacttggt 383
RESULT 12		
ASNPHYTAS		
LOCUS	ASNPHYTAS 2379 bp	DNA PLN 23-MAR-1994
DEFINITION	Aspergillus niger var awamori phytase gene, complete cds.	
ACCESSION	L02421	
VERSION	L02421.1	GI:166518
KEYWORDS	phytase.	
SOURCE	Aspergillus niger (strain Alk0243, sub-species awamori) DNA.	
ORGANISM	Aspergillus niger	
	Ascomycota; Pezizomycotina; Eurotiomycetes;	

Db 834 AACGTGACCACTTTGATGGAATAATATGCTTCTGAAGACATACAACTACAGCTTGGGT 893
 QY 384 gctgacgacttgactccattcggtgaaacacaaatggttaactctggtatttaattctac 443
 Db 894 GCAGATGACCTGACTCCCTTCGGAGACAGAGTAGTCAACTCCGGCATCAAGTTCTAC 953
 QY 444 agaagatacaagggttgggtgataaagattgttccattcttagagcttctctggtctgac 503
 Db 954 CAGCGATACGAATGCTCACAAGGAACATCATTCGTTTCATCCGATCCCTCTGGCTCCAG 1013
 QY 504 agagtattctctgctgctgaaagttcattgaaggtttccaaatctgctaaagttggtgac 563
 Db 1014 CGCTGTATCCCTTCGGCGAGAATTCATTTGAGGGCTTCAGAGCACCAGGCTGAAGGAT 1073
 QY 564 ccagggtgtaaccacacacaaagcttccagttattaaagttattattccagaaaggtgct 623
 Db 1074 CCTCGTGCAGCGCGGCAATGCTCGCCCAAGATCGAGCTGGTCAATTCAGAGGCCAGC 1133
 QY 624 ggttaacaaacacttgacacaggttctgttactgcttctgaagaattctgaattgggt 683
 Db 1134 TCATCCAAACAACTCTCGACCCAGGACCTGCACTGTCTTTGAAGACAGCGAATTGGCC 1193
 QY 684 gacagcgtgaagctaaactctgctgttttctgctccactattagagcttagattgaa 743
 Db 1194 GATACCGTCGAAGCCAAATTCACCGCCAGGTTCGCCCTCCATTCGTCACAGCTCGAG 1253
 QY 744 gctcacttgccaggtgttaacttgactgacgaagcgtgttgaacttgatggacatggt 803
 Db 1254 AACGACCTGCTGGGGTGACTCTACACACACAGAAAGTGAAGTCACTCTACGACATGTC 1313
 QY 804 ccattgacactgttctgtagaactctgacgctactgaactgaattgtctcaattctgacttg 863
 Db 1314 TCCTTCGACACCATCTCCACGACGACCTGACACCAAGCTGTGCCCTCTCTGTGACCTG 1373
 QY 864 ttcactacagcaaatgattcaatgaagactacttgaacttcttgggttaagactactcgt 923
 Db 1374 TTCACCCATGACGATGATGATCCATGACACTGACACTGACCTCCCTGAAAAAATACTACGC 1433
 QY 924 tacggtgctgtaacccatttggttccagctcaaggttgggttctggttaacgaattgatt 983
 Db 1434 CATGGCGCAGTAACCGCTCGGCCGACCCAGGGCGCTCGGCTACGCTAACGAGCTCATC 1493
 QY 984 gctagattgactactctccagttcaagaacacacttctactaacacacacttctgactct 1043
 Db 1494 GCCGCTCACCCACTGCGCTGTCCAGATGACACAGCTCCCAACACACACTTGGACTCG 1553
 QY 1044 aaccagctacttcccaattgaacgctacttctgacgactgaacttctctcacgacaact 1103
 Db 1554 AACCCAGCTACCTTCGGCTCACTCTACTCTCTACGCGGACTTTTCCACAGATAACGCG 1613
 QY 1104 atggtttctatttcttctgcttgggtttgtgacacaggttactaaagcattgttctact 1163
 Db 1614 ATCATCTATCTCTCTTGTGCTTGGTCTGTACACGGCCTTAAGCCGCTGTCTACACG 1673
 QY 1164 tctgttgaattctattgaagaactgacgctgactgcttcttctgactgttccatcgct 1223
 Db 1674 ACCGTGAGAAATATACCCAGACAGATGGGTCTCTCTCTGTGACGCTTCCGTTTGT 1733
 QY 1224 gctagagcttaagtgaatgatgaatggaagctgaaagaaacattggttagagtt 1283
 Db 1734 TCAGCTGTGTAGCTGTAGATGATGATGCCAGGCGGACGAGAGCGGCTGTGCTGCTGTC 1793
 QY 1284 ttggttaacgacagaggttcttccattgacggttctggttgggtgacaaagttgggttagat 1343
 Db 1794 TTGGTTAATGATCGGTTGTCCTCGCTGATGGGTGTCCAAATGATGCTTTGGGAGATGT 1853
 QY 1344 aagagagacgactctggttaaggttcttcttctgtagatctggtggttaacgtggaagaa 1403
 Db 1854 ACCCGGATAGCTTTGTGAGGGGTTGACTTTGCTAGATCTGGGGGTGATTGGCGGAG 1913
 QY 1404 ttttctgctta 1414
 Db 1914 TGTTCTGCTTA 1924

RESULT 13 ASPHYTASE

LOCUS
 DEFINITION

ACCESSION
 VERSION

KEYWORDS
 SOURCE

ORGANISM

REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE

REFERENCE
 AUTHORS

TITLE
 JOURNAL

FEATURES

source

exon
 CDS

sig_peptide
 intron
 exon
 mat_peptide

BASE COUNT
 ORIGIN

Query Match
 Best Local Similarity
 Matches

43.4%;
 65.3%;
 908;

Score 618.2; DB 13;
 Pred. No. 1.8e-153;
 Mismatches 483;

Indels 0; Gaps 0;

QY 24 gtcgtgctactgtccattgcccacttctgctggttccacatcccggtaccgcttgggtcct 83

Db 797 GTGGGACTACTGATCGCTGACTATCTGTGACAGATGATCCTCCGACTGGCAGTCCCGCC 856

QY 84 cgtgtaattctcactcttctgacactgttgacggtggtttaccaatgtttccagaatt 143

Db 857 TCAGAAATCAATCCAGTTCGATACGGTGCATCAGGGGTATCAATGCTCTCCGAGACT 916

QY 144 tctcaactgtgggttcaactacttccattctcttcttcttcttcttcttcttcttcttct 203

Db 917 TCAGATCTTTGGGGTCAATACGACCGGTCTTCTCTCTGCGCAACCAACGCGTCTATCTCC 976

QY 204 ccagacgttcccaagggtttagagttacttcttcttcttcttcttcttcttcttcttct 263

ASPHYTASE 2665 bp DNA PLN 27-APR-1993
 Aspergillus niger myo-inositol hexaphosphate phosphohydrolyase gene,
 complete cds.

M94550
 M94550.1 GI:166520

myo-inositol hexaphosphate phosphohydrolyase; phytase.
 Aspergillus niger (library: EMBL 3; NRRL 3135) DNA.

Aspergillus niger
 Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

1 (bases 1 to 2665)
 Mullanev,E.J., Gibson,D.M and Ullah,A.H.

Positive identification of a lambda gt11 clone containing a region of
 fungal phytase gene by immunoprobe and sequence verification

Appl. Microbiol. Biotechnol. 35, 611-614 (1991)
 92000601

2 (bases 1 to 2665)
 Mullanev,E.J.

Sequence of the Aspergillus niger (ficuum) phytase gene
 Unpublished (1992)

Location/Qualifiers
 1..2665

/organism="Aspergillus niger"
 /db_xref="taxon:5061"
 /tissue_lib="EMBL 3; NRRL 3135"

<683..726,829..2188)
 join(683..726,829..2188)
 /note="phytase gene"

/codon_start=1
 /protein_id="AAA32705.1"
 /db_xref="GI:166521"

/translation="MGVSAVLLPLLYLLSGVTSGLAVPSNRNOSCDTVDQYQCFSET
 SHLQYAPFFSLANESVISEVPAGCVTFQVLSRHGARYPTDSKGYKYSALIEEI
 QNATTFQKTAFLKTYNSLIGADDTLPFGQELVNSGKIFQYVESLNRNVPFIRS
 SGSSVIASGKKFTGQFQSTKLPRAQPGQSPKLDVISEASSNNTLDPGCTVF
 ESELDADVEANFTAFVPSIRQLNDLSGLTLDTEVYLMDCSFDTISTVDT
 KLSFPCDLTFHDEWINDYLOSLKLYGHGAGNPLGPTQGVGYANELIARLTHSPVHD
 DTSNHTLDSPATPLNLTLYADFSDNGILSILFALGLYNGKPLSTVVENITQT
 DGFSAATVPFASRLYVEMMCQAEPLVRLVNDRVVPLHGPCVDALGRCTRDSEV
 RGLSFARSGGDWAECA"

sig_peptide
 intron
 exon
 mat_peptide

727..828
 829..2188
 842..2176

/standard_name="phytase"
 /EC_number="3.1.3.8"
 /function="catalyzes the hydrolysis of phytic acid"
 /product="myo-inositol hexaphosphate phosphohydrolyase"

BASE COUNT 633 a 808 c 574 g 650 t

ORIGIN

Query Match
 Best Local Similarity
 Matches

43.4%;
 65.3%;
 908;

Score 618.2; DB 13;
 Pred. No. 1.8e-153;
 Mismatches 483;

Indels 0; Gaps 0;

QY 24 gtcgtgctactgtccattgcccacttctgctggttccacatcccggtaccgcttgggtcct 83

Db 797 GTGGGACTACTGATCGCTGACTATCTGTGACAGATGATCCTCCGACTGGCAGTCCCGCC 856

QY 84 cgtgtaattctcactcttctgacactgttgacggtggtttaccaatgtttccagaatt 143

Db 857 TCAGAAATCAATCCAGTTCGATACGGTGCATCAGGGGTATCAATGCTCTCCGAGACT 916

QY 144 tctcaactgtgggttcaactacttccattctcttcttcttcttcttcttcttcttcttct 203

Db 917 TCAGATCTTTGGGGTCAATACGACCGGTCTTCTCTCTGCGCAACCAACGCGTCTATCTCC 976

QY 204 ccagacgttcccaagggtttagagttacttcttcttcttcttcttcttcttcttcttct 263

Db	2057	TTGGTTAATGATCGCGTTCCTCCGCTGCATGGGTCTCCGGTTGATGCTTTGGGAGATGT	2116
Qy	1344	aagagagacgaattcgttgaaggtttcttcttcgtaagctggtgtaactggaagaa	1403
Db	2117	ACCCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGTGATGGCGGAG	2176
Qy	1404	tggttcgctta	1414
Db	2177	TGTTTTCCTTA	2187
RESULT 14			
LOCUS	AL9451	6756 bp	DNA
DEFINITION	phytase gene locus of plasmids pAF 2-3, pAF 2-6, pAF 2-7.		PAT 10-JUN-1994
ACCESSION	AL9451		
KEYWORDS	AL9451.1	GI:583193	
SOURCE	synthetic construct.		
ORGANISM	artificial sequence.		
REFERENCE	1 (bases 1 to 6756)		
AUTHORS	van Gorcom, R.F.M., van Hartingsveldt, W., van Paridon, P.A., Veenstra, A.E., Luiten, R.G.M. and Selten, G.C.M.		
TITLE	Cloning and expression of microbial phytase		
JOURNAL	Patent: EP 0420358-A 40 03-APR-1991;		
FEATURES	GIST-BROCADES N.V.		
source	Location/Qualifiers		
1..6756	/organism="synthetic construct"		
1..6756	/db_xref="taxon:32630"		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	join(210..254,357..1712)		
1..6756	/gene="phytase gene"		
1..6756	/gene="phytase gene"		
1..6756	/number=1		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	/codon_start=1		
1..6756	/transl_table=1		
1..6756	/protein_id="CAA01463.1"		
1..6756	/db_xref="GI:583194"		
1..6756	/translation="MGVSAVLLPLYLISGVTSLAVPASRNQSCDVTDOGYOCFSET SHLWQYAPFFSLANESVISEPVAGCRVTFQAQLSRHGRYPTDSKKYKSLIIEI QGNATVFDGKVFALKTNYSLGADLTFFGQELVNSGIRKIFORYESLIRNIVFIR SGSSRVIAAGKFFIEGQSTKIDKPRAPQGGSPKIDVISESSNNTLDPGCTVVF EDSELADVEANFTATFVPSIRORLENDLSGVTITDEVLYLMDMSFDTISTVDT KLSPCDLFTDHEWINDYLOSLKKYGHGAGNPLGTQGVYANELIARLTHSPVHD DTSSNHTLDSPATFPLNSTLYADSFHDNGIISILFALGLYNGTKPLSTTTVENITQT DGFSSAWTVFASRLYVEMQCAEQEPLRVLYNDVRVPLHGCPVDALGRCTRDSFV RGLSFARSGGDWAECA"		
1..6756	/number=1		
1..6756	357..1715		
1..6756	/gene="phytase gene"		
1..6756	/number=2		
1..6756	1712 c	1712 g	1714 t
1..6756	3 others		
BASE COUNT	1615 a	1712 c	1712 g
ORIGIN			
Query Match	43.4%	Score 618.2;	DB 9; Length 6756;
Best Local Similarity	65.3%	Pred. No. 1.9e-153;	
Matches	908;	Conservative	0; Mismatches 483; Indels 0; Gaps 0;
Qy	24	gtcgtgctactgtccattgccacctgttgcgttccacatcccggtaccgcttggtcct	83
Db	324	GTGGGACTACTGATCGCTGACTATCTGTGAGAGTACCTCCGGACATGGCAGTCCCGGC	383
Qy	84	cgtgtaattcactctgtgacacgttgacggttggttaccaattgtttcccaagaatt	143

Db	384	TCGAGAAATCAATCCAGTTCGGATACGGTCGATCAGGGGTATCAATGCTTCCGAGACT	443
Qy	144	tetcaacttggaggtcaataactcttcattctctcttggctgcgaatctgctattct	203
Db	444	TCGCATCTTTGGGGTCAATACGACCGGTCTCTCTGTGGCAACGAATCGGTCACTCC	503
Qy	204	ccagacgttccaaagggttgttagagtacttttgttcaaagtttgtctagacacggtgt	263
Db	504	CCTGAGGTGCCCGCGATGACAGTCACTTTCGTGAGGTCTCTCCGTCATGAGGCG	563
Qy	264	agatacccaacttctctaagtctaagaagtactctctgttgttgaagctattcaaaag	323
Db	564	CGGTATCCGACCGACTCCAAGGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG	623
Qy	324	aagctactgcttccaagggttaagttaagcttcttgaagctataactacactttgggt	383
Db	624	AAGCGACCACTTTGACGGAAATATGCGCTTCTGTAAGACATACAACTACAGCTTGGGT	683
Qy	384	gctgaacacttgactccattcgttgaaacaaatggttaactctggttattgaattctac	443
Db	684	GCAGATGACTGACTCCCTTCGGGAACAGAGGCTAGTCAACTCCGGCATCAAGTTCTAC	743
Qy	444	agaagatacaaggcttggcttagaagaattgttccattgtagagcttctggttctgac	503
Db	744	CAGCGGTACGATCGTCCACGAAGAACATCGTTCCATTATCCGATCTCTGGCTCCAGC	803
Qy	504	agagttattgctctgctgaaagttcattgaaggtttccaatctgctaaattggtctac	563
Db	804	CGCGTATCGGCTTCGGCAAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGGAT	863
Qy	564	ccagtgcttaaccacaacgaacttccgaattattaaagttattattccagaaggtgct	623
Db	864	CCTCTGCGCCAGCCGGCCAAATCGTCGCCCAAGATCGAGTGGTGTATTTCCGAGGCCAGC	923
Qy	624	ggttacaacaacactttggaccacggtttgtgtactgctttcgaagaatctgaaattgggt	683
Db	924	TCATCCAAACAACACTTCGACCCAGGCACCTGCACCTGTCTCGAAGACAGCGAATGGCC	983
Qy	684	gacgactgaagctaacttcaactgactgacgaagacgttgttaactgtatgagctagattgaa	743
Db	984	GATACCGTCGAAGCCAAATTTTACCGCCACGGTTCGTCCCTCCATTCGTCAACGCTGGAG	1043
Qy	744	gctcaattgacaggtgttaacttgactgacgaagacgttgttaactgtatgagctatggt	803
Db	1044	AACGACCTGTCCGGTGTGACTCTCAGACACAGAGTGACCTACCTCATGGACATGTGC	1103
Qy	804	ccattcaacactgtgctagaacttctgaactcaactcaattgtctccattctgtgacttg	863
Db	1104	TCCTTGCACACCAATCTCCACAGCACCGTCGACCAAGGTGTCCCCCTTCTGTGACCTG	1163
Qy	864	ttaactcacgaagatggattcaatacacactacttgaatttgggaagtactacggt	923
Db	1164	TTCAACCATGACGAATGGATCACTACGACCTACCTTCCAGTCTCTTGAAGAAAGTATTACGGC	1223
Qy	924	taagggtggttaaccattgggtccagctcaaggttgtgttctgttaacgaattgatt	983
Db	1224	CATGGTTCAGGTAAACCGCTCGGCCGACCCAGGCGTCGGCTACCGTACAGAGCTCATPC	1283
Qy	984	gctagattgactactctccaggttcaagacacacttctactaacacacttttgactct	1043
Db	1284	GCCGGTCTGACCCATCGCCTGTCCACGATGACACCAAGTTCCAAACACACTTTGGACTCG	1343
Qy	1044	aaccagactatttccattgaaagctacttttgaagctgaacttctcacgacaacact	1103
Db	1344	AGCCCGCTACCTTTCCGCTCAACTACTCTCTACGCGGACTTTTCGATCAACAGCGC	1403
Qy	1104	atggtttctatttcttcgcttgggtttgtaacaggtactaaagccatttgtctactact	1163
Db	1404	ATCATCTCAATCTCTTTCGCTTTAGGTCTGTACAACGGCACTAAAGCCGCTATCTACCAG	1463
Qy	1164	tctgttgaactctattgaagaaactgaaggttaagctgcttcttggactgttccattgct	1223

Db	1464	ACC GTG GGA AAT ATAC CAC CAG A CAG ATG GAT TCT CGT CTG CTT GCG ACC GTT CCG GTT GCT	1533
Qy	1224	gct agag ctt cao gtt ga aat at gca at gta ag act gaa ag gaa cca ttt ggt tag ag ttt	1283
Db	1524	TGC GTT TGT ACG TCG AG ATG ATG CAG GTT CAG GC GAG CAG CAG CCG TCG TCG GTC	1583
Qy	1284	tgg tta ac gac aca gagg ttg tcc at tgc cao ggt tgg tgg tgc aca ag ttg ggt tag at g	1343
Db	1584	TTG GTT AAT GAT CG GTT GTT CCG CTG CAT GGG TGT CCG GTT GAT GTT TGG GAG ATG	1643
Qy	1344	aag aca gac ag act cgt tga ag ttg tct ttc ogt ag at ctt ggt gta act ggg aaga a	1403
Db	1644	ACC CGG ATG ACT TTG TGA GGG GTT GAG CTT TGT AG ATCT GGG GTG AT TGG CGG AG	1703
Qy	1404	tgt ttc gct tta 1414	
Db	1704	TG TTT TGT TTA 1714	
RESULT 15			
IL13429			
LOCUS	IL13429	6756 bp	DNA
DEFINITION	Sequence 31 from patent US 5436156.	PAT	26-JUL-1995
ACCESSION	IL13429		
VERSION	IL13429.1	GI:910770	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 6756)		
AUTHORS	Van Gorcom, R.F.M., Van Hartingsveldt, W., Van Paridon, P.A., Veenstra, A.E., Luiten, R.G.M. and Seiten, G.C.M.		
TITLE	Cloning and expression of phytase from aspergillus		
JOURNAL	Patent: US 5436156-A 31 25-JUL-1995;		
FEATURES	Location/Qualifiers		
source	1..6756		
BASE COUNT	1615 a 1714 c 1712 g 1715 t		
ORIGIN	/organism="unknown"		
Query Match 43.4%; Score 618.2; DB 10; Length 6756;			
Best local Similarity 65.3%; Pred. No. 1.9e-153;			
Matches 908; Conservative 0; Mismatches 483; Indels 0; Gaps			
Qy	24	gtcgtgctactgtccattgccacctgtgtccacatccggtacgccttgggtctct	83
Db	324	GTGGCCTACTGATCGCTGCTATCTTGTGCAGAGTCACTCCGGACTGGCAGTCCCGGCC	383
Qy	84	cgtggaattctcaactctgtgacactgttgacggtgtgtaccaatgtttcccagaatt	143
Db	384	TCGAGAAATCAATCCAGTTGGATACGGTGCATCAGGGGTATCAATGCTTCCGAGACT	443
Qy	144	tctcaactgtgggtccaatactctccattctcttcttggctgacgaattcgtattct	203
Db	444	TCGCATCTTTGGGTCAATACGACCGCTTCTCTCTGTCACAAAGTCCGTCATCTCC	503
Qy	204	ccagacgttccaaagggtgttagagttacttttggttccaaagttttgttagacacggtgct	263
Db	504	CCTGAGTGGCCCGCGGATCGAGTGCATTTGCTGCTAGCTCTCTCCGTCATGGAGCG	563
Qy	264	agatacccaacttcttctaagttctaagaagtactctgtgttgatgaacttctaaag	323
Db	564	CGGTATCCGACCGACTCCAAGGCCAAGAACTCCGCTCTCATTTAGGAGATCCAGAG	623
Qy	324	aacgtactgttccaaagggtgaagtcgcttttttgaagacttacaactacactttgggt	383
Db	624	AACCGACCACTTTTGACGGAAATATGCTTCTCTGAAGACATACAACTACAGCTTGGCT	683
Qy	384	gctgacgacttgactccattcgggtgacacaaaatgttaactctggtattaaagttctac	443
Db	684	CGATGACCTGACTCCCTTCGAGAACAGGAGCTAGTCACTCCGCGCATCAAGTTTAC	743

Sat Oct 27 15:25:36 2001

Qy 444 agaagatacaaggcttggctagaagaattgttcattcgttagagcttctggttctgac 503
 Db 744 CAGCGGTACGAATCGCTCACAAGGAACATCGTTCATTCATCGATCTCTGGCTCCAGC 803
 Qy 504 agagtattgtctctgctgaagaattcattgaaggtttccaatctgctgaagtgtgac 563
 Db 804 CGCGTGTGCTCCGCGCAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGGAT 863
 Qy 564 ccaggtgttaaacaccacacagcttctccagttattaaagttattattccagaagtgct 623
 Db 864 CCTCGTCCAGCGCGGCAATCTGCCCAAGATCGAGCTGGTGTATTTCCGAGGCCAGC 923
 Qy 624 ggttacaacaacatttgaccacggtttgtgtactgttttcgaagaattctgaattgggt 683
 Db 924 TCATCCAAACAACCTCTCGACCCAGGACCTGCACATGTCTTCGAAGACAGCGAATTGGCC 983
 Qy 684 gacgagcttgaagcttaacttcaactgtgttttgcctccactattagagctagattggaa 743
 Db 984 GATACCGTCGAAGCAATTTCCACCGCCAGCTTCGTCCCTCCATTCGTCAACGCTGGAG 1043
 Qy 744 gctcacttgccaaggtgttaacttgaactgaagaacagcttgttaacttgatggacatgtg 803
 Db 1044 AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTGACCTACCTCATGGACATGTC 1103
 Qy 804 ccattcgaactgttgcctagaacttctgaacttgaacttcaacttctccacttctgacttg 863
 Db 1104 TCCCTGACACCATCTCCACGACGCGTCGACACCAAGCTGTCCCGCTTCTGTGACCTG 1163
 Qy 864 ttaactcaagcagaatgattcaatacactacttgaacttgaacttcttgggttaagtactacgt 923
 Db 1164 TTCACCATGACGAATGGATCAACTACGACTACCTTCCAGTCTTGAAGATATTACGGC 1223
 Qy 924 tacggtgctgttaaccattggttcagctcaaggttgggttcttgaacgaattgatt 983
 Db 1224 CATGGTGCAGGTAAACCGCTCGCGCCGACCCAGGCGCTGCGCTACGCTAACGAGCTCATC 1283
 Qy 984 gctagattgactcaacttccagttcaagaacacacacttctactaaccacacttggactct 1043
 Db 1284 GCCGCTGTGACCACTCGCTGTCCACGATGACACACAGTTCACACACACTTTGGACTCG 1343
 Qy 1044 aaccagctacttccattgaactgaacttcttgaacttcttcaagcagaacact 1103
 Db 1344 AGCCCGCTACCTTCCGCTCACTACTCTCTACGCGGACTTTTCGATGACAAACGGC 1403
 Qy 1104 atggttctatttcttctgcttgggttgaagaacagctgaacgaattgcttactact 1163
 Db 1404 ATCATCTCCATTCCTTTGCTTTAGGTCTGTACAGCGCACTAAGCCGCTATCTACCACG 1463
 Qy 1164 tctgttgaactctattgaagaacagctgaacttcttgaacttcttgaacttcttgaactcgt 1223
 Db 1464 ACCGTGGAGAAATACCCACAGCAGATGGATTCGCTCTGCTTGGACGGTTCGGTTTGT 1523
 Qy 1224 gctagagcttacgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1283
 Db 1524 TCGGCTTTGTACGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1583
 Qy 1284 tgggttaacagagagattgttccattcagcttgggttgaagaattggaagttagatgt 1343
 Db 1584 TTGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1643
 Qy 1344 aagagagacgacttctggtgaaggttcttctcgttagatctgtgtggaactgggaagaa 1403
 Db 1644 ACCGGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTTGGCGGAG 1703
 Qy 1404 tgttctgctta 1414
 Db 1704 TGTCTTGTCTTA 1714

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 16:39:50 ; Search time 2563.44 Seconds
(without alignments)
5258.469 Million cell updates/sec

Title: US-09-488-265-25

Perfect score: 1426

Sequence: 1 tatatgaattcatggcgctg.....ttcgcttaagaattcatata 1426

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

us-09-488-265-25.rst

Sat Oct 27 15:25:38 2001

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	151	10.6	802	227	AQ324539	mgxb0019C
2	117.8	8.3	226	AQ255459	mgxb0014M	
3	113.8	8.0	581	104	AJ274007	
4	88.6	6.2	753	225	AQ163004	mgxb0021D
5	81.8	5.7	699	225	AQ162040	mgxb0010F
6	79.4	5.6	769	225	AQ361474	mgxb0004I
7	76	5.3	810	227	AQ161556	mgxb0008K
8	72	5.0	669	227	AQ361495	mgxb0004K
9	45.2	3.2	1016	221	CNS04ESQ	
10	43	3.0	545	224	AQ140737	Tetraodon
11	42	2.9	468	151	BF637070	HS_3124_A
12	41.8	2.9	341	109	AV054552	BF637070 NF049F02L
13	41.8	2.9	481	225	AV054552	AV054552
14	41.2	2.9	231	25	AV129427	AQ0202816 RPC111-48
15	41.2	2.9	234	109	AV064132	AV129427
16	41.2	2.9	236	109	AV054718	AV064132
17	41.2	2.9	247	28	AV293643	AV054718
18	41.2	2.9	264	108	AV002759	AV293643
19	41.2	2.9	269	109	AV077122	AV002759
20	41.2	2.9	278	25	AV114552	AV077122
21	41.2	2.9	285	109	AV081956	AV114552
22	41.2	2.9	297	109	AV060181	AV081956
23	41.2	2.9	300	109	AV053884	AV060181
24	41.2	2.9	306	25	AV118366	AV053884
25	41.2	2.9	342	109	AV047696	AV118366
26	40.4	2.8	482	30	AV423243	AV047696
27	40.2	2.8	265	109	AV049962	AV423243
28	40.2	2.8	295	26	AV12676	AV049962
29	39.6	2.8	227	25	AV161297	AV12676
30	39.6	2.8	249	109	AV056003	AV161297
31	39.6	2.8	268	109	AV091633	AV056003
32	39.6	2.8	266	109	AV049564	AV091633
33	39.6	2.8	280	109	AV079904	AV049564
34	39.6	2.8	270	109	AV078019	AV079904
35	39.6	2.8	285	109	AV100697	AV078019
36	39.6	2.8	289	25	AV120593	AV100697
37	39.6	2.8	289	109	AV055109	AV120593
38	39.6	2.8	290	109	AV053894	AV055109
39	39.6	2.8	291	25	AV151459	AV053894
40	39.6	2.8	296	109	AV058485	AV151459
41	39.6	2.8	297	109	AV050799	AV058485
42	39.6	2.8	298	109	AV067677	AV050799
43	39.6	2.8	300	109	AV054136	AV067677
44	39.6	2.8	303	25	AV118481	AV054136
45	39.6	2.8	303	25	AV164971	AV118481

ALIGNMENTS

RESULT 1
 LOCUS AQ324539/c
 DEFINITION mgxb0019C01r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0019C01r, DNA sequence.
 ACCESSION AQ324539
 VERSION AQ324539.1 GI:4116391
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 802)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL COMMENT

Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 408.
 Location/Qualifiers
 1..802

FEATURES
 source

/organism="Magnaporthe grisea"
 /strain="70-13"
 /db_xref="taxon:148305"
 /clone="mgxb0019C01r"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH, Site_1: HindIII, Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."
 BASE COUNT 166 a 218 c 209 g 208 t 1 others
 ORIGIN

Query Match 10.6%; Score 151; DB 227; Length 802;
 Best Local Similarity 55.2%; Pred. No. 5.7e-33;
 Matches 342; Conservative 0; Mismatches 266; Indels 12; Gaps 2;

QY 755 aggtgttaactgactgacgaagacgtgtgttaactgtgacatgtgtccattgtgactgttcaactcacga 814
 DB 614 AGGGTGAACCTCAGCGCCAAAGCGGTGAGATCATGAGTCTGCCCGTTTGAGAC 555
 QY 815 tgtgtctagaactctgagctactcaattgtctcattgtgactgttcaactcacga 874
 DB 554 CGTGGCAGATCCNCAAGCCAC-----CCTGTGCGAGTCTGCACGCTGTTTACGCAACG 501
 QY 875 cgaatgattcaatacagactacttgcgaattcttgggttaagtactacggttgcgtgtctgg 934
 DB 500 AGACTGGGAGGATATGACTATCTCCAGACACTGGGGAAGTGTATGTTACGCAATGG 441
 QY 935 taaccattgggtccagctcaagggtgtgtgttcgttaacgaattgattgcttagattgac 994
 DB 440 CAACCCCTGGGCTCCACCAAGGGTGGGCTTCGTCAACGAGCTCATCGCGAGGCTGCT 381
 QY 995 tcactctcagttcaagaccacacttctactaaccacactttggacttaaccagctac 1054
 DB 380 CCAAAAGCCCGTTGAAGACCACACAAATACCAATCGAGCTCGACAGCGACCCATCGAC 321
 QY 1055 ttccattgaacgtactattgtacgtgacttctctcacgacacacactatggtttctat 1114
 DB 320 GTTCCCACTAGACAAAAAGCTGTACGCGGACTTTAGCCATGATAACGATATGCTGGCAT 261
 QY 1115 ttctctcgtttgtttgtacaacgttacttaacgaattgcttactactcttctgtgaac 1174
 DB 260 CTACGCCCGCTGGGATTTACACGCCACCGCCCGCGATTTCGGTCCCAAAAAAGGAGAG 201
 QY 1175 ta-----ttaagaataacgtacgtgttactcgtgtcttcttggactgttccattcgtctag 1228
 DB 200 AAGGAGCGCGAGGAGCTACGCGGTTCTCTCGTCAGTGGCGGTACCGTTCGAGCGAG 141
 QY 1229 agcttacgttgaatgatgcaatgtgaagctgaaaaaaccattggttagagtttgg 1288

Sat Oct 27 15:25:38 2001

```

Db 140 GATCTTTGTTGAAAAATGACTTGGCAGGCGCAGAACGAGGAGCTTGTGAGATCTCTGGT 81
QY 1289 taacgacagagtggttcattgacaggttggtggttgacaaagttggttagatgaagag 1348
Db 80 CAACGACAGGCTGAGCGCCCTGACCAACTGGCATGCCGACAGTAGTAGTGGTCTGTTGCAGCT 21
QY 1349 agacgacttcggtgaaggtt 1368
Db 20 GAGCAAGTTCGTTGAGAGCT 1

RESULT 2
LOCUS A0255459 837 bp DNA GSS 23-OCT-1998
DEFINITION mxg00014M15r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mxg00014M15r, DNA sequence.
ACCESSION A0255459
VERSION A0255459.1 GI:3779774
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 837)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 60
High quality sequence stop: 277.
FEATURES
Location/Qualifiers
1..837
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mxg00014M15r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 248 a 154 c 188 g 243 t 4 others
ORIGIN

Query Match 8.3%; Score 117.8; DB 226; Length 837;
Best Local Similarity 56.2%; Pred. No. 2.9e-23;
Matches 221; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 748 atttgccaggtgttaactgactgacgaagagcgttggttgaacttgatgacatgtgtccat 807
Db 140 ATTTTCAGGTGCTGGGTAAACCGGTCGAAGACGCTTCTAAATCTTATGGGACCTGTGCACCA 199
QY. 808 tcgacactgttgcagaaactctgcagcgtactcaattgtctcattctgtgactgttca 867

```

```

Db 200 TGGAAACACCGCGACTTTTGAAAAAAGCTTGGCAAACTTTCACCGCTTTCACAACTCTCTTTA 259
QY 868 ctacgacgaatgattgattacgaactactactactactactactactactactactactactacg 927
Db 260 CGGAAGCAGATTGGGTAATAATGATTGTTCTGTCACGCGTGAAGAAATGTTGATGACATGACG 319
QY 928 gtgctggttaacccatttggttcagctcaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 987
Db 320 GAACGGTAATCTTTGGGCCCACTATGGCGTGGGATGGTAAACGAACTATTATTCAC 379
QY 988 gattgactcactctccagttccaagaccacactcttacttaaccacacactttagacttaacc 1047
Db 380 GATTAAACCGAAGCCAGTCGCAAGATCAACCATGTCGCAATACGACACTTGCATGATGAC 439
QY 1048 cagctacttccattgaacgtactctgtgactgactgactgactgactgactgactgactgact 1107
Db 440 CGAAACCTTACTCTCTACAAAGCAATTTGTATGCTGATTTATCCATACGACGATATTA 499
QY 1108 ttctatttctctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Db 500 TAGGGATTACCGCGGCTTTGGGCTATTATTAACG 532

RESULT 3
LOCUS AJ274007 581 bp mRNA EST 29-DEC-1999
DEFINITION AJ274007 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae
cDNA clone Ma1628, mRNA sequence.
ACCESSION AJ274007
VERSION AJ274007.1 GI:6433380
KEYWORDS EST.
SOURCE Metarhizium anisopliae.
ORGANISM Metarhizium anisopliae
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Metarhizium.
REFERENCE 1 (bases 1 to 581)
AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.
TITLE EST analysis of the insect pathogenic fungus Metarhizium anisopliae
JOURNAL Unpublished (1999)
COMMENT Contact: Screen SE
Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA.
FEATURES
Location/Qualifiers
1..581
/organism="Metarhizium anisopliae"
/strain="ARSEF 2575"
/db_xref="taxon:5530"
/clone="Ma1628"
/clone_lib="Metarhizium anisopliae ARSEF 2575"
/note="Vector: Unizap; Metarhizium anisopliae was grown on
insect cuticle for 24 hours. A cDNA library was
constructed in the unidirectional Lambda vector, Unizap"
BASE COUNT 138 a 153 c 146 g 144 t
ORIGIN

Query Match 8.0%; Score 113.8; DB 104; Length 581;
Best Local Similarity 58.5%; Pred. No. 3.8e-22;
Matches 237; Conservative 0; Mismatches 162; Indels 6; Gaps 2;

QY 153 tgggggtcaatactctccattctcttcttggctgacgaatctgtctatttccagaggtt 212
Db 163 TGGGGCCACTACTCGCGGTACTCTC---AGCACACAGAGGATCTATTAGCCTTGATC 219
QY 213 ccaaaaggggttagaggttacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 272
Db 220 CCGTCGGGATGCGAGGTCACTACCTTTGCTTCAATCTTCTCGACATGGTCCCGCAACCG 279
QY 273 acttcttcttaagtcttaagaagtactctgtctgttattgattgaagtatttcaaaagagctact 332
Db 280 ACGCCCGGCAAGTCCCAAGCCCTACAGAGGACCTGTGTGAGCGGATCCAAAAGACGTCAAG 339

```



```

QY 333 gcttcaagggttaagtacgtttcttgaagacttaactacactttgggtgctgaacac 392
Db 340 GACTATGGGAGGCGCTTCGAGTCTTCTCAAGATTACAAATACATCTAGTCTTGACGAC 399
QY 393 ttgactccattcgttgacacaaatggttaactctggtattgaattctacagaagatac 452
Db 400 CTCACCCCTTGGCGAGCAGGAATGTCAGTCTCGAAAAAGATTCCTCAAGCCGCTAT 459
QY 453 aaggcttgctagaaa---gattgtccattcgttagagcttctggttctgacagatt 509
Db 460 CAGAAGTTAGCGGAGGATTCGACCCATCCCTTTGTCGAGCCCTCGGCTCTGAAAGATC 519
QY 510 attgcttctgctgaagagtttcaaggtttccaatctgctaaag 554
Db 520 GTCATGTGACGCGCAAGATTGTTGTCATGGCTTCTACAAAGCCCAAG 564

RESULT 4
AQ163004 753 bp DNA GSS 09-SEP-1998
LOCUS clone mgxb0021D19r, CUGI Rice Blast BAC Library Magnaporthe grisea genomic
DEFINITION clone mgxb0021D19r, DNA sequence.
ACCESSION AQ163004
VERSION AQ163004.1 GI:3559405
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 753)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 448.
FEATURES
source
Location/Qualifiers
1. 753
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0021D19r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 208 a 180 c 156 g 206 t 3 others
ORIGIN

Query Match 6.2%; Score 88.6; DB 225; Length 753;
Best Local Similarity 51.3%; Pred. No. 9.4e-15;
Matches 230; Conservative 0; Mismatches 212; Indels 6; Gaps 1;

```

```

QY 101 ttgtgactgtttgacggttggttaccaatgtttccagaaaatttctcacttgtgggtca 160
Db 189 TTTTGTATGGCGCAGCGCGGGTTTCAATTTAAAGCAGCAATCACTATAATTGGGGCCA 248
QY 161 atactctccattctctcttctgttgacgaatctgctatttctccagacgtttccaa---- 216
Db 249 GTATGCACCACTTTTCGTTGCACCGCTCAGCACCATCGGATTATGATTCTCTGCTACCTT 308
QY 217 --aggttgtagttactttctggttcaagtttcttagacacggtgtagatacccaac 274
Db 309 GCCAGGGTGTAGTACTAGTTTGCCCAAGTCTTTCAGAGGCACGGGGCTCGATACCCCAAC 368
QY 275 ttcttctaagtcataagaagtagtctgctgttgaagctattcaaaagaacgctactgc 334
Db 369 CGCTCAAAACCGGGCGGAAATTCGACACAGATTCCACCGACTCCAAACCTCAGCTACCGG 428
QY 335 ttcaagggttaagtacgctttcttgaagaacttaactacactttgggtgctgacgactt 394
Db 429 TAGTGGCGTATTGAATACTACTATTAATAAATACTAATAATACAAATCTCGGAGTCGAGGAATN 488
QY 395 gactccattcgttgacacaaatggttaactctggttataagttctacagaagatacaa 454
Db 489 GAATGATTTCGGCGCGCTCGGCAACAGAGAACTCCGGGTATTNATTTTACCAAGATACCA 548
QY 455 ggccttggtagaagaattgttccattcgttagagcttctggttctgacagagttattgc 514
Db 549 AAATCTGCCAGAGAGGAGCAACCACTTTATTCGCTATGATCACAANAACCGGTGTTGA 608
QY 515 ttctgctgaagaagttcattgaaggtttc 542
Db 609 TAGCGCAGAACTCTGGGCTAGGGGTTTC 636

RESULT 5
AQ162040/c 699 bp DNA GSS 09-SEP-1998
LOCUS clone mgxb0010F06r, CUGI Rice Blast BAC Library Magnaporthe grisea genomic
DEFINITION clone mgxb0010F06r, DNA sequence.
ACCESSION AQ162040
VERSION AQ162040.1 GI:3558441
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 699)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 285.
FEATURES
source
Location/Qualifiers
1. 699
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0010F06r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with

```

of rice world wide. It is a filamentous ascomycete with a haploid genome ($n=7$) of approximately 40 Mbp. Rice blast is an important fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters

BASE COUNT	211 a	185 c	157 g	213 t	3 others
ORIGIN					

	Query Match	5.6%;	Score 79.4;	DB 227;	Length 769;
	Best Local Similarity	51.7%;	Pred. NO. 4.6e-12;		
	Matches 233;	Conservative	0;	Mismatches 211;	Indels 7;
					Gaps
Qy	98	ctctgtgacactgttgcagtgagttaccacaaattctccacagaaaattctccacttgttgggg	157		
Db	273	CTGTGTTTCATGCGCCAGCCGCGGTTTTCAAATTAAAGCAGCAATCACTCATAAATTTGGG	332		
Qy	158	tcaataactctccattctctctcttctgttgcagcaatctgctattctccagacggtccaa-	216		
Db	333	CCAGTATGCACCATTTTTCGTTGCACCGTCAGACCAATCGGATTATGATTCCTCTGCTAC	392		
Qy	217	-----agggtgtgagtagtaacttcctcgaattttgtctagacacggtgctagatacc	271		
Db	393	CTTGCCAGGGTGTAGCATATACGTTTGCCCAAGTCTTTCAGAGGCACGGGCTCGATACCC	452		
Qy	272	aactctcttaagtctaagaagtagtactctgctttgattgaagctattcaaaagaacgctac	331		
Db	453	ATCCCTCAATACCCGGCGCGAAATTTCTCGAACACAGATTCACCGACTCCAAACCTCAGCTAC	512		

513 CGGTAGTGGCGTATTGCAATAACTACATTAAAAAATACAAAATCTCGAGCTCAAGGA 572

392 cttgaactccattcgtgtgaacaacaaatggttaactcttggtattaagttctacagaagata 451

573 ATTGAATGATTTGGGCGCTGCCAAACAGAGAAC-CTCGGGTATATTTTACCAAGAATA 631

452 caagcgttgggtgagaagaattggtccattcgttagagctctcgtggttcttcacagaggttat 511

632 CCAAAATCGCCAGAAGAACCACTTTATTCGTCTGATGACAAAAAACCGCTGTT 691

512 tgctctcgtgaaaagtctcattgaaggttc 542

692 TGATAGCGCAGAACCTTGGGCTTAGGGTTTC 722

RESULT	7
LOCUS	AQ161556/c
DEFINITION	810 bp . DNA
	GSS
	09-SEP-1998
	BAC Library Magnaporthe grisea genom
	clone mxgb0008K02r, DNA sequence.
ACCSSION	AQ161556
VERSION	AQ161556
KEYWORDS	GI:3557957
SOURCE	GSS.
ORGANISM	Magnaporthe grisea. Magnaporthe grisea Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae. 1 (bases 1 to 810) Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A. A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome Unpublished (1998)
REFERENCE	Contact: Dean RA
AUTHORS	Clemson University Genomics Institute
TITLE	Clemson University
JOURNAL	100 Jordan Hall, Clemson University, Clemson, SC 29634
COMMENT	Tel: 864 656 5737

Fax: 864 656 4293
Email: rdean@clmson.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 321.
Location/Qualifiers
1. .810

FEATURES

source
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0008K02"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request." 207 t

BASE COUNT

196 a 219 c 188 g 207 t

Query Match

Best Local Similarity 5.3%; Score 76; DB 225; Length 810;
Matches 133; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1190 cgggtacgtgcttctggactgtccattcgctgctagagcttacgttgaaatgatgca 1249
Db 650 CGGTTCTGTCGCCGCGGGCGGTACGTTCCGACGCGAGGATGTTGTTGAAAAAATGAC 591
QY 1250 atgtgaagctgaaaggaaccattggttagagtttgggttaagacagagattgttccatt 1309
Db 590 TTGCGCAGGGCAGAACGAGGAGCTGTGTGAGAACTCTGTCACGACGAGGCTGACGCCGT 531
QY 1310 qcacgggtgtgtgtgacaggttggttagatgataagagacagacttcgtgaaggttt 1369
Db 530 GCACAACTGGATGCCGACAGTAGTGGTCTGTCACGCTGAGCAAGTCTGTGAGAGCTT 471
QY 1370 gtcttgcgtagatctggtgtgaactggggaagaatgtttcgccttaaga 1417
Db 470 AAGCTCTGCGAGGAGTGAGGTCGCTGGGATCAATGTTTGTCTGAGA 423

RESULT 8

AQ361495/c
LOCUS 669 bp DNA GSS 03-FEB-1999
DEFINITION mgxb0004K10r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0004K10r, DNA sequence.
ACCESSION AQ361495
VERSION AQ361495.1 GI:4211334
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea

REFERENCE 1 (bases 1 to 669)
AUTHORS Yu, X., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737

Fax: 864 656 4293
Email: rdean@clmson.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 465.
Location/Qualifiers
1. .669

FEATURES

source
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0004K10r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request." 233 t

BASE COUNT

185 a 119 c 131 g 233 t

Query Match

Best Local Similarity 5.0%; Score 72; DB 227; Length 669;
Matches 288; Conservative 0; Mismatches 321; Indels 14; Gaps 1;

QY 458 ttggttagaaagatttcttcattcgtagagcttgggtttgacagagattgttc 517
Db 627 TTGCCCAAGATAAATAAATTTTATCCCTTCTAATGGAGCGAGTGTGTTAAAG 568
QY 518 tgcgtgaaagtcattgaaggtttcccaatctgtaagtgtggtgacccaggtgtaaccc 577
Db 567 CGCTTAAATGGCTGACGGCTTTCATCAAACTCTTTAACCCGATAAACTCGAATCA 508
QY 578 acaccaagcttctccagttatttaacgttattattccagaaggtgctggttacaacacac 637
Db 507 ACCGCATCAATTTCCCGCTTTTCAATTTTAAATTTCTGTGCGAAGCAATATAACACAC 448
QY 638 ttgggaccacgggttggtagctgcttcgaagaatcgaattgggtgacgacgttgaagc 697
Db 447 TTTAAATTAATCTCGCTTGCACCAAAATACGACAAATCCAAAAGCTCAAAAATTTGCAAAA 388
QY 698 t-----aacttcactgctgttttcgctccacctattagagctagattgaa 743
Db 387 ACTACCAAAAAAAGGGGAAANAATTTTGACGGAAACCGCAACCGTTGAAT 328
QY 744 gctcactgcccaggtgttaactgactgaogaagacgttgttaactgtatggacatgtgt 803
Db 327 CAAAACCTTACCAGGGGCAAAATTTATCAGTTTACGAAATTTCTCAACTTATGGAACATGT 268
QY 804 ccattcgacactgttctagaacttctgcagctactcaattgtctccattctgtgacttg 863
Db 267 GGTATGGAAACAACGGCTAATATTGTAAAAACGGGTATTTTATCGCCCATTTGCAATATT 208
QY 864 ttcactcacgacgaatggattcaatcagactacttgcactctttgggttaagactacggt 923
Db 207 TTTGACAAACAAGATTTGGGAGCGCTTTGACGGCTTTTCAATAAGTAAATTAATGTTATACC 148
QY 924 tacggtgctggtaacccattgggtccagctcaaggtgtgtgttcgttaacgaattgatt 983
Db 147 CACTTTTATGGTAACCCCTTTAGGGCCGTCATTGGGAGTAGGTTGGGTAATGAATGATATC 88
QY 984 gctagattgactcactctccagttcaagaccacacttctactaacacacacttggactct 1043
Db 87 GCGAGGTTATTACGAAGACCTGTACAAGATTCCACCAGCACCAACCAACCAATTAAGTCTGC 28

clone 1810031C08, mRNA sequence.

ACCESSION AV054552
VERSION AV054552.1 GI:5154299
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 341)
AUTHORS Akañkara, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C.,
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara,
Y., Suzuki, H., Suzuki, H., Tateono, M., Tomaru, Y., Tomimaga, N.,
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
Muranatsu, M., Okazaki, Y. and Hayashizaki, Y.
TITLE
RIKEN Mouse ESTs
UNPUBLISHED (1999)
CONTACT: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998)) Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES	Location/Qualifiers
source	i. .341
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="1810031C08"
	/clone_lib="Mus musculus pancreas C57BL/6J adult"
	/sex="male"
	/tissue_type="pancreas"
	/dev_stage="adult"
BASE COUNT	86 a 60 c 101 g 94 t
ORIGIN	
Query Match	2.9%; Score 41.8; DB 109; Length 341;
Best Local Similarity	52.6%; Pred. No. 0.34;
Matches 91; Conservative	0; Mismatches 82; Indels 0; Gaps 0;

```

FEATURES      further details.
source        Location/Qualifiers
1. .341
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /clone="1810031C08"
   /clone_lib="Mus musculus pancreas C57BL/6J adult"
   /sex="male"
   /tissue_type="pancreas"
   /dev_stage="adult"
BASE COUNT   86 a . 60 C 101 g 94 t
ORIGIN
Query Match      2.9% ; Score 41.8 ; DB 109 ; Length 341 ;
Best Local Similarity 52.6% ; Pred. No. 0.34 ;
Matches 91 ; Conservative 0 ; Mismatches 82 ; Indels 0 ; Gaps 0 ;

```

```

source
i: 341
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="1810031C08"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
BASE COUNT      86 a      60 c      101 g      94 t
ORIGIN

Query Match      2.9% ; Score 41.8; DB 109; Length 341;
Best Local Similarity 52.6% ; Pred. No. 0.34;
Matches 91; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```

```

BASE COUNT      86 a      .60 c      101 g      94 t
ORIGIN

Query Match      2.9%;      Score 41.8;      DB 109;      Length 341;
Best Local Similarity 52.6%;      Pred. No. 0.34;
Matches 91;      Conservative 0;      Mismatches 82;      Indels 0;      Gaps 0;

```

Query Match	2.9%	Score 41.8;	DB 109;	Length 341;
Best Local Similarity	52.6%;	Pred. No. 0.34;		
Matches 91;	Conservative	0;	Mismatches	82;
			Indels	0;
			Gaps	0;

QY 1166 tgttgaatctattgaagaaactgacggtttagctgttctcttggactgttccattcgtgc 1225

Db 151 TGGTGGGCCGTTCCAGCAGCTGGAGTTGGCCCTGCTGGTGTGTGTTCCATCCACTGC 210

QY 1226 tagagcttacgttgaaatgatgcaatgtgaagctgaaaaggaaccattggttagagttt 1285

Db 211 CGCGCCCCAGTTGAGGAGAGCAAGAGTGGAGGCAAGAGGAAGATCCGAGGAGCTTGA 270

QY 1286 ggtaacgacagagttgttccattgcacggttggtgttgacagattgggta 1338

D_b 271 AGATGACATGGGCTTTGTTTTTGGACTAAACTGTTTTTGAAGTTGGCTA 323

RESULT 13
A0202816

LOCUS	AQ202816	481 bp	DNA	GSS	20-APR-1999
DEFINITION	RC111-48J14	TK RPCR-11	Homo sapiens	genomic clone	RC111-48J14

DEFINITION RPEC11-48J14.TK RPEC1-11 Homo sapiens genomic clone RPEC1-11-48J14,
DNA sequence.
ACCESSION AC202816

KVVHQBDS
 CSS
 VERSION AQ202816.1 GI:3613607
 ACCESSION AQ202816

KEYWORDS	SOURCE
GSS.	human.

AUTHORS

JOURNAL,

INTRODUCTION

3-1-1 Vondra; Boujuba

Fax: 81-298-36-9098

Thermostabilization and thermoact.

source

/organism

`/db xref="taxon:10090"`

```
clone lib="Mus muscul
```

[illegible]

```
new_stage="ADDIT"
```

420 129 023

Rest Local

607

0673 cott

43 CTGC

1225 ctaq.

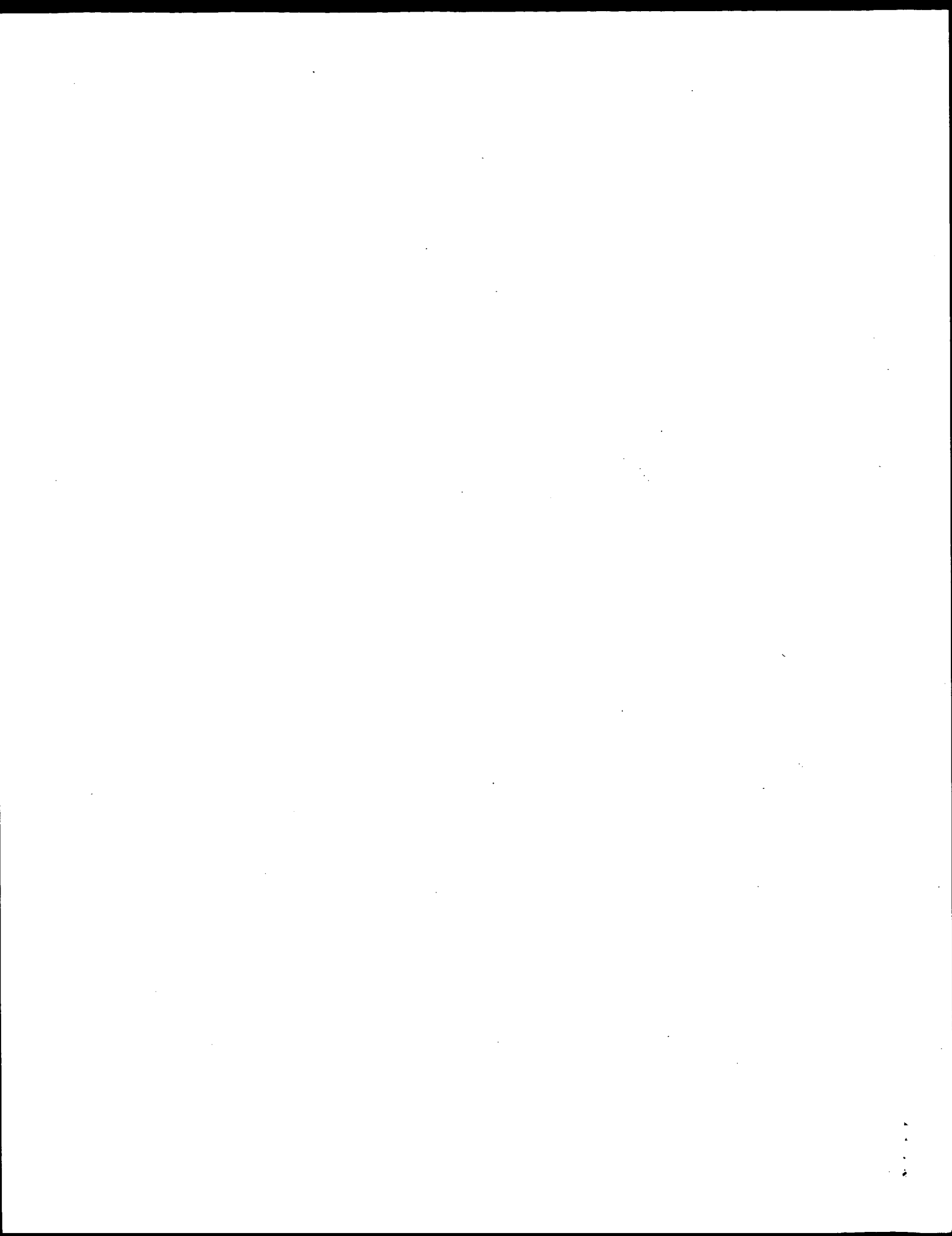
103 CCGO

1285 tgggt

163 AAGH

Job time: 6790 sec

[illegible]



Db 241 ORLENDLSGVLTDTTEVYLYMDMCSFTDITSTVDTKLSPCDLFTDHEWLHYDYLQSLK 300
QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYHGAGNPLGTPGVGYANELIARLTHSPVDDTSSNHTLDSNPATFPLNATLYADFS 360
QY 361 HDNTMYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAMTVPFASRLYVEMMQCEAKEP 420
QY 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVRLGSLFARSFGDWAECFA 467
Db 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVRLGSLFARSFGDWAECFA 467
RESULT 2
JN0656
3-Phytase (EC 3.1.3.8) A precursor - Aspergillus niger
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phytA protein
C:Species: Aspergillus niger
C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0656; S28456
R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.
A:M.J.J.
Gene 127, 87-94, 1993
A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)
A:Reference number: JN0656; MUID:9322284
A:Accession: JN0656
A:Molecule type: DNA
A:Residues: 1-467 <VAN>
A:Cross-references: GB:216414; NID:q2392; PIDN:CAA78904.1; PID:g2393
A:Experimental source: strain NREL3135
A:Note: parts of the sequence, including the amino end of the mature protein, were confirmed
C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and inosic acid
C:Genetics:
A:Gene: phyA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-467/Product: 3-phytase A #status experimental <MAT>
F:27,59,105,120,207,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #sta
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 75.4%; Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 1.7e-141;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
QY 1 MGVFVLLSTATLFGSTGALGRGNSHSDTVDDGYQCFPEISHLWGQYSPFFSLADE 60
Db 1 MGVSALLPLXLSGVTSLGLAVPASRQSSCDIVDQGYQCFSETSHLWGQYAPFFSLANE 60
QY 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSKYSALIEAIOKNAFAKGYAFKLTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSKYSALIEAIOKNAFAKGYAFKLTYN 120
QY 121 YLGLADDTLTPGEQMVNSGKIFVRRYKALARKTVPPVRSASDRVTASAEKIEGFQSA 180
Db 121 YSLGADDTLTPGEQELVNSGKIFVRRYKALARKTVPPVRSASDRVTASAEKIEGFQSA 180
QY 181 KLADPGANPHQASPVNVIPEGAGYNNTLDHGLCTAFEESELGDVVEANFTAVFAPP 240
Db 181 KLKOPRAQPGQSKIDVWVISEASSNNTLDPGCTVFEDSELADTVANFTAVFAPP 240
QY 241 ARLEAHLPGVNLTDVVDVMDMCPFDVARTSDATQSLSPCDLFTDHEWTDYDYLQSLG 300
Db 241 ORLENDLSGVLTDTTEVYLYMDMCSFTDITSTVDTKLSPCDLFTDHEWTDYDYLQSLK 300
QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYHGAGNPLGTPGVGYANELIARLTHSPVDDTSSNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNTMYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAMTVPFASRLYVEMMQCEAKEP 420
QY 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVRLGSLFARSFGDWAECFA 467
Db 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVRLGSLFARSFGDWAECFA 467
RESULT 3
JN0482
3-Phytase (EC 3.1.3.8) A - Aspergillus ficum
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phytA protein
C:Species: Aspergillus ficum
C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
C:Accession: JN0482; PN0023
R:Ullah, A.H.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 192, 747-753, 1993
A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemi
A:Reference number: JN0482; MUID:93249451
A:Accession: JN0482
A:Molecule type: protein
A:Residues: 1-441 <ULL>
A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 178, 45-53, 1991
A:Title: Cyclohexanediol modification of arginine at the active site of Aspergillus
A:Reference number: PN0023; MUID:91298982
A:Accession: PN0023
A:Molecule type: protein
A:Residues: 48-70 <ULL>
C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydrate (Asn) (covalent) #sta
F:58,338/Active site: Arg, His #status predicted
F:59/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 72.9%; Score 1801; DB 1; Length 441;
Best Local Similarity 76.0%; Pred. No. 1.3e-136;
Matches 333; Conservative 42; Mismatches 63; Indels 0; Gaps 0;
QY 27 NSHSCDVTVDGYQCFPEISHLWGQYSPFFSLADESAISPOVPGKRCRVTFVQVLSRHGARY 86
Db 4 NOSSCDTVDDGYQCFSETSHLWGQYAPFFSLANESVISEPVPAGCRVTFVQVLSRHGARY 63
QY 87 PTSSKSKYSALIEAIOKNAFAKGYAFKLTYNITLGGADDTLTPGEQMVNSGKIFYRR 146
Db 64 PDSKSKYSALIEAIOKNAFAKGYAFKLTYNITLGGADDTLTPGEQELVNSGKIFYQR 123
QY 147 YKALARKTVPPVRSASDRVTASAEKIEGFQSAKLADPGANPHQASPVNVIPEGAGY 206
Db 124 YESLTRNIVPTFRRSSGSRVITASGKKIEGFQSTKLKDPRAQPGQSSPKIDVWVISEASS 183
QY 207 NNTLDHGLCTAFEESELGDVVEANFTAVFAPPFARLEAHLPGVNLTDVVDVMDMCPFF 266
Db 184 NNTLDPGCTVFEDSELADTVANFTAVFAPFARLEAHLPGVNLTDVVDVMDMCPFF 243
QY 267 DTAVARTSDATQSLSPCDLFTDHEWTDYDYLQSLGKYYGYGAGNPLGPAQGVFNELIAR 326
Db 244 DTISTSTVDTKLSPCDLFTDHEWTDYDYLQSLGKYYGYGAGNPLGPAQGVFNELIAR 303
QY 327 LPHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTMYSIFFALGLYNGTKPLSTTSV 386
Db 304 LTHSPVDDTSSNHTLDSNPATFPLNATLYADFSHDNGLIISILFALGLYNGTKPLSTTV 363
QY 387 ESTEETDGYAASWTVPFAARAYVEMMQCEAKEPILVRVLNDRVPLHGGVVDALGRCTR 446
Db 364 ENITQDGFSSAMTVPFASRLYVEMMQCEAKEPILVRVLNDRVPLHGGVVDALGRCTR 423
QY 447 DDFVEGLSFARSFGNWE 464

Db 424 DSFVRLSFASSGGDAE 441

RESULT

PABYC

acid phosphatase (EC 3.1.3.2) repressible, precursor - yeast (Saccharomyces cerevisiae)
 N;Alternate names: acid phosphatase PHO5; protein YBR0814; protein YBR093C
 C;Species: Saccharomyces cerevisiae
 C;Date: 19-Feb-1984 #sequence-revision 30-Sep-1991 #text-change 12-Nov-1999
 C;Accession: S05795; A38792; S48260; S45961; A00777; A38793; S41855; B25241; A25367; A27
 R;Bajwa, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
 Nucleic Acids Res. 12, 7721-7739, 1984
 A;Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeast
 A;Reference number: S05794; MUID:85037940
 A;Accession: S05795
 A;Molecule type: DNA
 A;Residues: 1-467 <BAJ>
 A;Cross-references: EMBL:X01079; NID:g4162; PIDN:CAA25555.1; PID:g758282
 A;Note: The authors translated the codon TAC for residue 272 as Thr
 A;Accession: A38792
 A;Molecule type: protein
 A;Residues: 18-45 <BAJ2>
 R;Mannhaupt, G.; Stucka, R.; Ehnl, S.; Vetter, I.; Feldmann, H.
 Yeast 10, 1363-1381, 1994
 A;Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
 A;Reference number: S48255; MUID:95208357
 A;Accession: S48260
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-467 <MAN>
 A;Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55598.1; PID:g476051
 A;Note: The nucleotide sequence was submitted to the EMBL Data Library, April 1994
 R;Feldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
 submitted to the Protein Sequence Database, August 1994
 A;Reference number: S45927
 A;Accession: S45961
 A;Molecule type: DNA
 A;Residues: 1-467 <FE2>
 A;Cross-references: EMBL:Z35962; NID:g536364; PIDN:CAA85046.1; PID:g536365; GSPDB:GN0000
 R;Arima, K.; Oshima, T.; Kubota, I.; Nakamura, N.; Mizunaga, T.; Toh-e, A.
 Nucleic Acids Res. 11, 1657-1672, 1983
 A;Title: The nucleotide sequence of the yeast PHO5 gene: a putative precursor of repress
 A;Reference number: A00777; MUID:83168913
 A;Accession: A00777
 A;Molecule type: DNA
 A;Residues: 1-35, 'Y', 37-129, 'G', 131-293, 'Q', 295-445, 'V', 447-461, 'DT', 464-465, 'K', 467 <AR
 A;Cross-references: EMBL:V01320; NID:g4158; PIDN:CAA24630.1; PID:g4159
 A;Accession: A38793
 A;Molecule type: protein
 A;Residues: 18-26, 'X', 28 <ARI2>
 R;Meyhack, B.; Bajwa, W.; Rudolph, H.; Hinnen, A.
 EMBO J. 1, 675-680, 1982
 A;Title: Two yeast acid phosphatase structural genes are the result of a tandem duplicat
 A;Reference number: S41855; MUID:84236032
 A;Accession: S41855
 A;Molecule type: DNA
 A;Residues: 1-30, 'T', 32-51, 'S', 53-75 <MEY>
 A;Cross-references: EMBL:M24178; NID:g172156; PIDN:AAA34868.1; PID:g172157
 R;Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G
 Mol. Cell. Biol. 6, 1855-1865, 1986
 A;Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
 A;Reference number: A93074; MUID:87064474
 A;Accession: B25241
 A;Molecule type: DNA
 A;Residues: 1-44 <TAI>
 R;Bergman, L.W.
 Mol. Cell. Biol. 6, 2298-2304, 1986
 A;Title: A DNA fragment containing the upstream activator sequence determines nucleosome
 A;Reference number: A25367; MUID:87064526
 A;Accession: A25367
 A;Molecule type: DNA
 A;Residues: 1-2, 'Y', 4-43, 'T', 45-51 <BER>

R;Silve, S.; Monod, M.; Hinnen, A.; Haguenauer-Tsapis, R.
 Mol. Cell. Biol. 7, 3306-3314, 1987
 A;Title: The yeast acid phosphatase can enter the secretory pathway without its N-ter
 A;Reference number: A27774; MUID:88038886
 A;Accession: A27774

A;Molecule type: DNA

A;Residues: 1-51, 'S', 53-60 <SIL>

A;Cross-references: GB:M17306

C;Genetics:

A;Gene: SGD:PHO5; MIPS:YBR093C

A;Cross-references: SGD:S0000297; MIPS:YBR093C

A;Map position: 2R

A;Note: YBR093C

C;Superfamily: yeast acid phosphatase

C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
 F;1-17/Domains: signal sequence #status predicted <SIG>

F;18-467/Product: acid phosphatase, repressible #status experimental <MAT>

F;75/Active site: His (phosphohistidine intermediate) #status predicted

F;97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (CO

F;337/Active site: His #status predicted

Query Match 16.3%; Score 403; DB 1; Length 467;

Best Local Similarity 26.2%; Pred. No. 1.8e-24;

Matches 126; Conservative 61; Mismatches 216; Indels 78; Gaps 15;

QY 6 VLLSTATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGQYSPFFSLADESAISP 65

DB 5 VVYSILAASLANAGTI--PLGLADVDVIGTKQKDFPEL---GGAGPYSPFGDYGISR 58

QY 66 DVPKGRVTEVQVLSRHGARYPTSSKKYSALIEAIQKNATAPKGAFLK-TYNYTLG 124

DB 59 DLPECEMKQLQMGVGRGERIPTVSLAKTIKSTYKLSNYTRQFNGLSFLNDDYEFFIR 118

QY 125 ADD-----LTPF-GEQMVNSGKIFVRRYKALKARKIVPF-VRASGSDRVIA 168

DB 119 DDDLEMETTFANSDVNLNPTGEMNAKRHARDFLAQYGVVENQTSFAVTSNSKRCHD 178

QY 169 SAEFIEGFSQAKLADPGANPHQASPVINVI-----PEGAGYNNLTHGLCTAFEESEL 223

DB 179 TQAFYFDG-----LGDQ-----FNITLQTVSEASAGANTLSACNSCPAW-DYDA 222

QY 224 GDDVEANFTAVFAPPFIRARLEAHLPGVNLDTEDVVNLMDMCPEDIVARTSATQISPCD 283

DB 223 NDDIVNEYDTYLLDDIAKRLNKKGLNLTSTDALETSWCAFEV-----NAKGYSDVCD 277

QY 284 LFTHDEWTDYDLSGLKYYGYGAGNPLGPAQGVGVFNELIARLTHSPVQDHTSTNHTLD 343

DB 278 IFTKDELVHYSYQDLHTYHEGPGYDIKSVGNLNFASVKLLKQSEIQDQ----- 329

QY 344 SNPATFPLNATLYADFSDNTWVSIFPALGLYNGTKPLSTTSVESIETDGYAASWTVPF 403

DB 330 -----KWLSTFHTDIDLFLTAGIIDDKNLTAEYVPMFGNT--FHRSWYVPQ 377

QY 404 AARAYVEMMOCEAEKPELVRVLVNDVRVPLHCGVKGCKGRDFF-----VEGLSFA 456

DB 378 GARVYTEKFCQ--SNDYFVRVINDAVVPIETCTGPGFSCIEINDFYDAEKRVAGTDFL 435

QY 457 R 457

DB 436 K 436

RESULT

S53476

acid phosphatase (EC 3.1.3.2) precursor - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein YAR071w

C;Species: Saccharomyces cerevisiae

C;Date: 05-May-1995 #sequence-revision 01-Sep-1995 #text-change 20-Apr-2000

C;Accession: S53476; JCI018

R;Bussey, H.; Keng, T.; Storms, R.K.; Vo, D.; Zhong, W.; Fortin, N.; Barton, A.B.; Ka

submitted to the EMBL Data Library, February 1994

A;Description: Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of th

A:Reference number: S53458

A:Accession: S53476

A:Molecule type: DNA

A:Residues: 1-467 <BUS>

A:Cross-references: EMBL:128920; NID:g1616966; PIDN:AAC09508.1; PID:g456155; MIPS:YAR071

R:Chen, J.Y.; Gong, Y.; Ao, S.Z.

Acta Biochim. Biophys. Sin. 21, 437-444, 1989

A:Title: The primary structure of acid phosphatase gene PHO11 in *S. cerevisiae* and compa

A:Reference number: JCI018

A:Accession: JCI018

A:Molecule type: DNA

A:Residues: 1-16, 'L', 18-149, 'H', 151-353, 'O', 355-422, 'G', 424-467 <CHE>

A:Note: This paper is in Chinese, with an English abstract

C:Genetics:

A:Gene: SGD:PHO11

A:Cross-references: SGD:S0000094; MIPS:YAR071w

A:Map position: 1R

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydroly

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-467/Product: acid phosphatase #status predicted <MAT>

F:74/Active site: Arg #status predicted

F:75/Active site: His (phosphohistidine intermediate) #status predicted

F:97,162,192,250,315,356,439,445,461/Binding site: carbohydrate (Asn) (covalent) #status

Query Match

Best Local Similarity 16.2%; Score 399; DB 2; Length 467;

Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCFPEISHLMQYSPFFSLADESAISPDVPGKRVTFVQVLSRHG 83

DB 21 PLGKLSIDIKIGTQTEIFPFL---GGSPYISFPDGYGSRDLPSCEMKQVQVGRHG 76

QY 84 ARYPTSSKKYKYSALLIATOKNATAFKGYAFK-----TNYNLT-----GADDL 128

DB 77 ERYPTVSKAKSIMTTWYKLSNYTQFSGALSFNDDEFFIRDTKMLETTILANSVNVL 136

QY 129 TPE-GEOQVNSGKIKYRYKALARKIVPF-VRASGSDRVIAAEKFIQEGFQSAKLADPG 186

DB 137 NPYTGEMNAKRHARDLAQGYGVNENQTSFAVTSNSNRCHDTAQYFIDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFESSELGDDVEANFTAVFAPPIRA 241

DB 191 -----FNISLQITSEASAGANTSAHSHSCPAWDD-DVNDLILKYDYTKYLSGIK 240

QY 242 RLEAHLPGVNLTDVVDVNLMDMCPDVTARTSDATQSPCDLFTHEWTOYDYLSLQSLG 301

DB 241 RLKKNKGLNLTSSDANTFFAWCAYENAR-----GYSIDICNIFTKDELVRFSYQDLET 295

QY 302 YGYGAGNPLGPAQGVGVNLIARTLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSE 361

DB 296 YYQTGPGYDVRVSRGANLFNASVKLKESEVQDQ-----KWLSEFTH 337

QY 362 DNTMVSIFFGALYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOQEAKEPL 421

DB 338 DTDILNLTITIGIDDKNNLTAEHVPFMENT--FHRSWYVPOGARVYTEKFOC--SNDTY 393

QY 422 VRVLVNDRVVPLHGGCGVDKLGCRKDDF-----VEGLSFAR 457

DB 394 VRYVINDAVVPIETCTGPGFSCSEINDFYAEKRVAGTDFLK 436

RESULT 6

S48996

N:Alternate names: PHO12 - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999

A:Accession: S48996; S59659

R:Macri, C.

submitted to the EMBL Data Library, February 1994

A:Description: The sequence of *S. cerevisiae* cosmid 9177.

A:Reference number: S46671

A:Accession: S48996

A:Molecule type: DNA

A:Residues: 1-467 <MAC>

A:Cross-references: EMBL:U00029; NID:g551322; PIDN:AAB69729.1; PID:g458917; MIPS:YHR2

R:Xu, L.

submitted to the EMBL Data Library, January 1995

A:Reference number: S59658

A:Accession: S59659

A:Molecule type: DNA

A:Residues: 1-16, 'L', 18-81, 'AR', 84-149, 'H', 151-467 <XUL>

A:Cross-references: EMBL:U19789; NID:g847754; PIDN:AAA73479.1; PID:g847755

C:Genetics:

A:Gene: SGD:PHO12

A:Cross-references: SGD:S0001258; MIPS:YHR215w

A:Map position: 8R

C:Superfamily: yeast acid phosphatase

C:Keywords: phosphoric monoester hydrolase

Query Match

Best Local Similarity 16.2%; Score 399; DB 2; Length 467;

Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCFPEISHLMQYSPFFSLADESAISPDVPGKRVTFVQVLSRHG 83

DB 21 PLGKLSIDIKIGTQTEIFPFL---GGSPYISFPDGYGSRDLPSCEMKQVQVGRHG 76

QY 84 ARYPTSSKKYKYSALLIATOKNATAFKGYAFK-----TNYNLT-----GADDL 128

DB 77 ERYPTVSKAKSIMTTWYKLSNYTQFSGALSFNDDEFFIRDTKMLETTILANSVNVL 136

QY 129 TPE-GEOQVNSGKIKYRYKALARKIVPF-VRASGSDRVIAAEKFIQEGFQSAKLADPG 186

DB 137 NPYTGEMNAKRHARDLAQGYGVNENQTSFAVTSNSNRCHDTAQYFIDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFESSELGDDVEANFTAVFAPPIRA 241

DB 191 -----FNISLQITSEASAGANTSAHSHSCPAWDD-DVNDLILKYDYTKYLSGIK 240

QY 242 RLEAHLPGVNLTDVVDVNLMDMCPDVTARTSDATQSPCDLFTHEWTOYDYLSLQSLG 301

DB 241 RLKKNKGLNLTSSDANTFFAWCAYENAR-----GYSIDICNIFTKDELVRFSYQDLET 295

QY 302 YGYGAGNPLGPAQGVGVNLIARTLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSE 361

DB 296 YYQTGPGYDVRVSRGANLFNASVKLKESEVQDQ-----KWLSEFTH 337

QY 362 DNTMVSIFFGALYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOQEAKEPL 421

DB 338 DTDILNLTITIGIDDKNNLTAEHVPFMENT--FHRSWYVPOGARVYTEKFOC--SNDTY 393

QY 422 VRVLVNDRVVPLHGGCGVDKLGCRKDDF-----VEGLSFAR 457

DB 394 VRYVINDAVVPIETCTGPGFSCSEINDFYAEKRVAGTDFLK 436

RESULT 7

PABYCC

N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c

C:Species: Saccharomyces cerevisiae

C:Date: 30-Sep-1991 #sequence_revision 09-Sep-1994 #text_change 05-Nov-1999

A:Accession: S48259; S45960; S05794; A25241; S44674

R:Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.

Yeast 10, 1363-1381, 1994

A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.

A:Reference number: S48255; MUID:95208357

A:Accession: S48259

A:Status: nucleic acid

A:Molecule type: DNA

A:Residues: 1-467 <MAN>

A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55597.1; PID:g476050

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
 R:Feidmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45927
 A:Accession: S45960
 A:Molecule type: DNA
 A:Residues: 1-467 <FE2>
 A:Cross-references: EMBL:Z35961; NID:q536362; PIDN:CAA85045.1; PID:q536363; GSPDB:GN0000
 R:Baiba, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
 Nucleic Acids Res. 12, 7721-7739, 1984
 A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeas
 A:Reference number: S05794; MUID:85037940
 A:Accession: S05794
 A:Molecule type: DNA
 A:Residues: 1-218, 'MKT', 222-467 <BAJ1>
 A:Cross-references: EMBL:X01080; NID:q4148; PIDN:CAA23557.1; PID:g9758281
 A:Note: The authors translated the codon AAT for residue 134 as Asp and TAC for residue
 R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G.
 Mol. Cell. Biol. 6, 1855-1865, 1986
 A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
 A:Reference number: A93074; MUID:87064474
 A:Accession: A25241
 A:Molecule type: DNA
 A:Residues: 1-44 <TAI>
 C:Genetics:
 A:Gene: SGP:PHO3; MIPS:YBR092C
 A:Cross-references: SGD:S0000296; MIPS:YBR092C
 A:Map position: 2R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoric monoester hydrola
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-467/Product: acid phosphatase, constitutive #status predicted <MAT>
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (coval
 F:337/Active site: His #status predicted

Query Match 16.0%; Score 394; DB 1; Length 467;
 Best Local Similarity 27.0%; Pred. No. 9.7e-24;
 Matches 117; Conservative 62; Mismatches 191; Indels 64; Gaps 13;
 QY 49 GOYSPFFSLADESAISDPVKGCRVTFOVLSRHGARYPTSSKSKYSALIEAIQKNATA 108
 Db 42 GGAGPVFSPGCDGISDLPEGCEMQLQMLARHGERIPYTSKAGIMKTWYKLSNYTRQ 101
 QY 109 PKGYAFLEK-TYNYTLGADD-----LTPF-GEQQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFLNDYDEFFIRDDDLMEETTFSNDVNLNPTTGEMDAKRHAREFLAQGYMFE 161
 QY 153 KIVPFF-VRASGSDRVIASAEKFELEGFQSALADPGANPHQASPVINVIIEGAGYNTLD 211
 Db 162 NQTSFFIFFAASSERVHDTAQYFIDGL-----GDQFNIQLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFEESELGDVEANFTAVFAPPFIRARLEAHLPGVNLTDVQVNLMDCPEDTVA 270
 Db 211 AGNACPCWDE-DANDILDKYDTYLLDILAKLENKENGLNLSKDNLTFAWCAYELNA 269
 QY 271 RYSDATQLSPFCDLFTHDWIOYDIYQSLGKYGYGAGNPLGPAQGVFVNLIELIALTHS 330
 Db 270 R-----GYSDVCDIETEDELVYSYGQDLVSFYQDGPYDMIRSVGANLFNATLKLKQS 324
 QY 331 PVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFFALGLYNGTGPLSTTSVESIE 390
 Db 325 ETQD-----LKVLSFTHTDITLNYLTAGIIDDKNLNTAEYVPFG 366
 QY 391 ETDGAASTVTPFAARAYVEMMOCEAEKPLVRLVNLVWVPLHGCYVDKLGCRKDDP- 449
 Db 367 NT--FHKSWYVPOGARVYTEKFCQ--SNDTYVRYVINDAVVETCTSGPGFSCEINDEY 422
 QY 450 -----VEGLSFAR 457
 Db 423 DYAEARVAGTDFLK 436

RESULT 8

S52495
 acid phosphatase homolog YDL024C - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D2815
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
 C:Accession: S52495; S67556
 R:Andre, B.; Vissers, S.; Urrestarazu, L.
 submitted to the EMBL Data Library, February 1995
 A:Description: The sequence of a 42 kb segment located on the left arm of chromosome
 A:Reference number: S52492
 A:Accession: S52495
 A:Molecule type: DNA
 A:Residues: 1-468 <AND>
 A:Cross-references: EMBL:Z48432; NID:9683669; PIDN:CAA88335.1; PID:9683673
 A:Experimental source: strain S288C
 R:Urrestarazu, L.A.; Andre, B.; Vissers, S.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67335
 A:Accession: S67556
 A:Molecule type: DNA
 A:Residues: 1-468 <URR>
 A:Cross-references: EMBL:Z74072; NID:q1430996; PIDN:CAA98583.1; PID:q1430997; MIPS:YD
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 4L
 C:Superfamily: yeast acid phosphatase

Query Match 15.0%; Score 370; DB 2; Length 468;
 Best Local Similarity 25.1%; Pred. No. 8.2e-22;
 Matches 116; Conservative 62; Mismatches 174; Indels 110; Gaps 14;

QY 49 GOYSPFFSLADESAISDPVKGCRVTFOVLSRHGARYPTSSKSKYSALIEAIQKNATA 108
 Db 43 GGSAPYFSPFANYGIPDIDPEGCHLTQVMIGRGEYPTRSEAKDIFVWYKISNTYTKG 102
 QY 109 PKGYAFLEK-TYNYTL-----GADDLTPF-GEQQMVNSGKIFRYRYKALAR 152
 Db 103 YEGSLSLFNGYEFFIPDESLEMETTLQNSIDVNLNPTTGEMNAKRHAREFLAKYKILME 162
 QY 153 KIVPFF-VRASGSDRVIASAEKFELEGFQSALADPGANPHQASPVINVIIEGAGYNTLD 211
 Db 163 NCTNFPIFTNSKRIYDTAQYFAEL-----GDGFNISLQ 197
 QY 212 HGCTAFEESELG-----DDVEANFTAVFAPPFIRARLEAHLPGVNL 253
 Db 198 ----TLSENSSSGANTLAAKSSCPNNSNANNILMSYSDYLENLSDRLENKGLNLS 253
 QY 254 DEDVNLMDCPEDTVARISDATQLSPFCDLFTHDWIOYDIYQSLGKYGYGAGNPLGP 313
 Db 254 RKDAALFSCAFEL-----NAKYSNICDIFSAAELIHYSETDLTSTFYQNGPGYKLIK 308
 QY 314 AQGVGFVN---ELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFF 370
 Db 309 SIGANLFNATVKLIRSAH-----LDOKVWLSFTHTDILNYLT 347
 QY 371 ALGLYNGTGPLSTTSVESIEETDGYAASVTVPFAARAYVEMMOCEAEKPLVRLVNDVR 430
 Db 348 TAGLIDDTNLNTNHPFFRDHS--YHRSWVTPQGARVYTEKFCQ--SNDYSYRVYVNDV 403
 QY 431 VPLHGCYVDKLGCRKDDFE-----GLSFARSQGNNEEC 465
 Db 404 VPIESCSSGPGFSCEGTTEYAYAKDLRGVSP-----YEDC 439

RESULT 9

JN0715
 3-phytase (EC 3.1.3.8) B precursor - Aspergillus ficum
 N:Alternate names: pH 2.5-optimum acid phosphatase
 C:Species: Aspergillus ficum
 C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast *Pichia pastoris*: Characterization
A:Reference number: J4285; MUID:96001238
A:Accession: J4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:g881955; PIDN:AAA85503.1; PID:g881956
A:Experimental source: GS115
C:Genetics:
A:Gene: pho1
A:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:8/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:345/Active site: His #status predicted

Query Match 14.2%; Score 350; DB 2; Length 468;
Best Local Similarity 25.4%; Pred. No. 3.3e-20;
Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;

QY 6 VLLSIATL---FGSTGALPGRNSHSCDTVDGGYOCFPEISHLMGOYSPF--FSLADE 60
Db 10 IILALATLQSVFAVELQHLVGLVNDPRPQRT-DQYVNLRLHGLG-GPYLYGNGWIAAE 67
QY 61 SAISPDVPGKGRVTFVOVLSRHGARYPTSSKSKYSALIE-AIQKNATAFKGYAFLKY 119
Db 68 SEI-----ESTIQAHLLMRHGERYSTNVGKLEALYKLLDADVEVPTGPLSPFDY 122
QY 120 NYTLG---ADLLTPFGGQMVNSGIRP---YRYKAL-----ARKIVFFVRASGDR 165
Db 123 DYFVSDAAWYEQETTKGYSGLTAFDFTLRERYDHLINTSEEGKL--SVWAGSQER 180
QY 166 VIASKEFIEGFSQAKLADPGANPHQASPVINVIPE---GAGYNNLTLDHGLCT----A 217
Db 181 VDTAKYFAOGFMKSNYTD-----WVEVVALEEEKSOGSLNLTARISCPNYNSHI 230
QY 218 FEESLGDV---EANTFAVFAPPFIRAEHLPGVNLTDVYVNLMDMCPFFDVARTSD 274
Db 231 YKGDFFNDIAERAD-----RLNTLSPGNITADDDIPTIALYCOFELNVGR- 278
QY 275 ATQLSPFDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAOGVGVNELIARLTHSPVD 334
Db 279 ----SSFCVLSREALTYAYLRDLGWYVNGNPNLGTIGYVIAN----- 321
QY 335 HTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPFAALYNGTK-PLSTTSVESIEETD 393
Db 322 --ATQLLENTAD-PRDYPLYFSFSDTDLLOVFTSLGLFNVDLPL-----DQIQFT 373
QY 394 GYAASWTVPPFAARAYVEMQC--EAEKEPLVRVLVNDVRVPLHGGVGDVKLGRKRDDEVE 451
Db 374 SFKSTEIVPMGARLLTERLLCTVEGEKYYVYRILNDVAFPLSDCSPGPGFSCPLNDYVS 433
QY 452 GLSPARSGGNWEE 464
Db 434 RLEALNEDSDFAE 446

RESULT 11
JN0890
acid phosphatase (EC 3.1.3.2) precursor - *Aspergillus awamori*
C:Species: *Aspergillus awamori*
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
A:Accession: JN0890
R:Pidington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0890
A:Molecule type: DNA
A:Residues: 1-479 <PID>

Query Match 14.2%; Score 350.5; DB 1; Length 479;
Best Local Similarity 27.9%; Pred. No. 3.1e-20;
Matches 117; Conservative 53; Mismatches 159; Indels 91; Gaps 16;

QY 49 GOYSPFSLADESAISPDVPGKGRVTFVOVLSRHGARYPTSSKSKY-SALIEAIOKNAT 107
Db 53 GPYSERVSY---GIADPTGCEVDQVIMVKHGERYPSAGKSIIEALAKVYSINT 108
QY 108 AFKGYAFLKTYNYTL-----GADLLT-PF-GEQMVNSGIRYRYKAL--ARKIVFF 157
Db 109 EYKGLDAFLNDWTYVYVNECYNAETTSYAGLLDAYNHGNDYKARYGHLWNGETVWVPF 168
QY 158 VRASGSDRVIASAEKLEGFQSAKLADPGANPHQASPVINVIPEGAGYNNLTLDHGLCTA 217
Db 169 F-SSGYGRVIETARKGEGF-----FGYNYSTNAALNII 201
QY 218 FEESLGL-----DDVEANFTAVFAPPFIRAEHLPGVNLTDVYVNLMD 262
Db 202 SESEVMGADSLTPTCDTNDQTTCDNLTYQLPQFKVAAARLNSONPGMNLTSADVYNLMV 261
QY 263 MCPEDTVARTSDATQLSPFDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAOGVGFVNE 322
Db 262 MASFELNAR----PFSNWINAFTQDEWVSFGYVEDLNYIYACGPGDKNMAAVGAVTANA 316
QY 323 LIARLTHSPVDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPFAALYNGTK-NGTKP 380
Db 317 SLTLNQG-----KEAGP-----LFFNLADHTNITPILAAALGVLPNEDLP 358
QY 381 LSTTSVESIEETDGYAASWTVPPFAARAYVEMQCEA-----EKEPLVRVLVNDVRVPLHGC 436
Db 359 L-----DRVAFGNPYSIGNIVPMGGHLLTIERLSQATALSQATYVRLVNLNEALPFPNC 413

RESULT 10
J4285
acid phosphatase (EC 3.1.3.2) precursor - yeast (*Pichia pastoris*)
N:Alternate names: orthophosphoric monoester phosphohydrolase (acid optimum); Pho
C:Species: *Pichia pastoris*
C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999
A:Accession: J4285
R:Payne, W.E.; Gannon, P.M.; Kaiser, C.A.

Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast *Pichia pastoris*: Characterization
A:Reference number: J4285; MUID:96001238
A:Accession: J4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:g881955; PIDN:AAA85503.1; PID:g881956
A:Experimental source: GS115
C:Genetics:
A:Gene: pho1
A:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:8/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:345/Active site: His #status predicted

Query Match 14.2%; Score 350; DB 2; Length 468;
Best Local Similarity 25.4%; Pred. No. 3.3e-20;
Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;

QY 6 VLLSIATL---FGSTGALPGRNSHSCDTVDGGYOCFPEISHLMGOYSPF--FSLADE 60
Db 10 IILALATLQSVFAVELQHLVGLVNDPRPQRT-DQYVNLRLHGLG-GPYLYGNGWIAAE 67
QY 61 SAISPDVPGKGRVTFVOVLSRHGARYPTSSKSKYSALIE-AIQKNATAFKGYAFLKY 119
Db 68 SEI-----ESTIQAHLLMRHGERYSTNVGKLEALYKLLDADVEVPTGPLSPFDY 122
QY 120 NYTLG---ADLLTPFGGQMVNSGIRP---YRYKAL-----ARKIVFFVRASGDR 165
Db 123 DYFVSDAAWYEQETTKGYSGLTAFDFTLRERYDHLINTSEEGKL--SVWAGSQER 180
QY 166 VIASKEFIEGFSQAKLADPGANPHQASPVINVIPE---GAGYNNLTLDHGLCT----A 217
Db 181 VDTAKYFAOGFMKSNYTD-----WVEVVALEEEKSOGSLNLTARISCPNYNSHI 230
QY 218 FEESLGDV---EANTFAVFAPPFIRAEHLPGVNLTDVYVNLMDMCPFFDVARTSD 274
Db 231 YKGDFFNDIAERAD-----RLNTLSPGNITADDDIPTIALYCOFELNVGR- 278
QY 275 ATQLSPFDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAOGVGVNELIARLTHSPVD 334
Db 279 ----SSFCVLSREALTYAYLRDLGWYVNGNPNLGTIGYVIAN----- 321
QY 335 HTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPFAALYNGTK-PLSTTSVESIEETD 393
Db 322 --ATQLLENTAD-PRDYPLYFSFSDTDLLOVFTSLGLFNVDLPL-----DQIQFT 373
QY 394 GYAASWTVPPFAARAYVEMQC--EAEKEPLVRVLVNDVRVPLHGGVGDVKLGRKRDDEVE 451
Db 374 SFKSTEIVPMGARLLTERLLCTVEGEKYYVYRILNDVAFPLSDCSPGPGFSCPLNDYVS 433
QY 452 GLSPARSGGNWEE 464
Db 434 RLEALNEDSDFAE 446

RESULT 11
JN0890
acid phosphatase (EC 3.1.3.2) precursor - *Aspergillus awamori*
C:Species: *Aspergillus awamori*
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
A:Accession: JN0890
R:Pidington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0890
A:Molecule type: DNA
A:Residues: 1-479 <PID>

A:Cross-references: GB:L02420; NID:g166481; PIDN:AAAL6897.1; PID:g166482
 A:Experimental source: strain ALK0243
 C:Comment: The highly similar enzyme from A. ficum has been shown to have 3-phytase (EC 3.1.3.2)
 A:Gene: aph

A:Introns: 261/1; 300/2; 335/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phospho-
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-479/Product: 3-phytase #status predicted <MAT>
 F:81.337/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted
 F:106,191,227,250,315,340,425,442,458/Binding site: carboxylate (Asn) (covalent) #status

Query Match 14.0%; Score 346.5; DB 1; Length 479;
 Best Local Similarity 27.4%; Pred. No. 6.5e-20;
 Matches 115; Conservative 54; Mismatches 160; Indels 91; Gaps 15;

QY 49 GOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKY-SALIEAIQKAT 107
 Db 53 GPYSERSVY-----GIARDPPTSCVDQVIMVKRGERYPSPSAGKDEEALAKVISINT 108
 QY 108 AFKCKYAFKTYNTL-----GADDT-PP-GEQOMVSGIKFYRYKAL--ARKIVPE 157
 Db 109 EYKGDLAFLNDWTYVPEYNECYNAETTSVPYAGLLDAYNHGNDYKARYGHLWNGETVVPF 168
 QY 158 VRASGRDVIASAEKFTIEGQSAKLADPGANPHQASPVINVIPEGAGYNTLDHGLCTA 217
 Db 169 F-SSGGRVETARKFEGP-----FGYNYSTNAALNII 201
 QY 218 FESELG-----DDVEANFTAVFAPPTR---ARLEAHLPGVNLTDDEVNLM 262
 Db 202 SESEVMGADSLTPTCDTNDQTTCDNLTYQLPQFKVAARLNSQNGMLTASDVYLMV 261
 QY 263 MCPDPTVARTSDATQSPFCDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVFVNE 322
 Db 262 MASFELNAR-----PFSNWINAFQDEWVSFGYVEDLNYIYICAGPGDKNMAAVGAVYANA 316
 QY 323 LIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLY--NGTKP 380
 Db 317 SLTLAAGPKPE-----AGSLFFNFADHTNTPTILALGVLIPNEDLP 358
 QY 381 LSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEA-----EKEPLRVLVNDRVPLHGC 436
 Db 359 L-----DRVAFGNPYSIGNIVPMGGHLLTIERLSQATALSDEGTYVRLVNEAVLPFND 413

RESULT 12
 A25326
 acid phosphatase (EC 3.1.3.2) precursor [similarity] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: A25326; T50405
 R:Elliot, S.; Chang, C.; Schweingruber, M.E.; Schaller, J.; Rickli, E.E.; Carbon, J.
 J. Biol. Chem. 261, 2936-2941, 1986
 A:Title: Isolation and characterization of the structural gene for secreted acid phosphatase from Schizosaccharomyces pombe
 A:Reference number: A25326; MUID:86140050
 A:Accession: A25326
 A:Molecule type: DNA
 A:Residues: 1-453 <ELL>
 A:Cross-references: GB:M11857; NID:g173422; PIDN:AAA5321.1; PID:g173423
 R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 2000
 A:Reference number: Z25068
 A:Accession: T50405
 A:Molecule type: DNA
 A:Residues: 1-453 <RIE>
 A:Cross-references: EMBL:AL137099; PIDN:CAB68657.1; GSPDB:GN000067; SPDB:SPBP4G3.02
 A:Experimental source: strain 972h(-); clone p1 p4G3
 C:Genetics:
 A:Gene: pho1; SPDB:SPBP4G3.02
 A:Map position: 2

Query Match 12.2%; Score 301; DB 2; Length 463;
 Best Local Similarity 25.4%; Pred. No. 2.7e-16;
 Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

C:Superfamily: yeast acid phosphatase
 C:Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
 F:68/Active site: Arg #status predicted
 F:69/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 13.2%; Score 326; DB 1; Length 453;
 Best Local Similarity 24.8%; Pred. No. 2.6e-18;
 Matches 103; Conservative 63; Mismatches 201; Indels 48; Gaps 8;

QY 35 DGGYOCFPEISHLGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKK 94
 Db 27 DCKDFKHEHLTSRSPYHKPYF-----YGPSIDFPTTCKIKQVHTLQRHGSRNPTGGNAF 81
 QY 95 YSALIEAIQK-----NATAFKGYAFLKTYNTL---GADDTLTFGEQOMVNSGK 142
 Db 82 DAVGLANFQORLLNGSVDPIDYSVGNPLSPFTPTVTEAANADALSSSRVLEFDMGRQ 141
 QY 143 FYRYKALARKIVFPVRASGSDRVIAAEKFTIEGQSAKLADPGANPHQASPVINVIPE 202
 Db 142 FYERYHELFASTYNIYTAQQRVVDSALWYGYGMF-----CEDVHNTNYILVSENA 194
 QY 203 GAGYNNTLDHCLCTAFEESELGDDVEANFTAVFAPPTRARLEAHLPGVNLTDDEVNLM 262
 Db 195 TAGNSLSSYNACPASDADFTTPALEAWRNVMYPPIRQLRNPFYSNYLNTDNDILNLYG 254
 QY 263 MCPDPTVARTSDATQSPFCDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVFVNE 322
 Db 255 ICSVEIALQ-----DYSEFCKLNSVDFLNEFYEGDLSFSYGMNSVKWGSIFGAVANS 309
 QY 323 LIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKPLS 382
 Db 310 LANSL--RSVENNTQ-----QVFFAETHDANIIPVETALGFFTDNTPEN 351
 QY 383 TTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEEKEPLRVLVNDRVPLHGC 437
 Db 352 PLTYSYQVHSHSKASEFVFAGNLITELFCEDSKY-YVRHLVNEEVFPLSDCG 405

RESULT 13
 S14119
 acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
 C:Accession: S14119; T40455
 R:Yang, J.; Schweingruber, M.E.
 Curr. Genet. 18, 269-272, 1990
 A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schizosaccharomyces pombe
 A:Reference number: S14119; MUID:91064763
 A:Accession: S14119
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <YAN>
 A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA40258.1; PID:g5007
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z21931
 A:Accession: T40455
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-463 <LYN>
 A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN000067; SPDB:SPBC428.03c
 A:Experimental source: strain 972h; cosmid c428
 C:Genetics:
 A:Gene: SPBC428.03c
 A:Map position: 2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 12.2%; Score 301; DB 2; Length 463;
 Best Local Similarity 25.4%; Pred. No. 2.7e-16;
 Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

Query Match 12.2%; Score 301; DB 2; Length 463;
 Best Local Similarity 25.4%; Pred. No. 2.7e-16;
 Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

Db 269 AIKDH-----SDPCSITPTSEFLNFEYDSDLQOAYGGGPVSEWASTLGCAYINNLAIDL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFALGLYNGTKPLSTTSVES 388
 Db 323 -----RNVNTPDFDRK-----VFLAFTHDSNIIPIVEAALGFPPDITFONPLPTDK 367
 QY 389 IEETDGYAASWTVPFAARAYVEMMOCEAEKPLRVLRVNDVVRVPLHGGCGVDKLGK- - - - C 444
 Db 368 NIYYSOKTSFVPFAGNLITELFFC-SDSKYYVRHLVNOQVYPLDLCGYPGSGTSDGLC 426
 QY 445 KRDDFVEGLSFARSGGN 461
 Db 427 ELQAYLNSPIRANSTSN 443

RESULT 15
 JE0369
 histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
 C:Accession: JE0369
 R: Mullane, E.J.; Ullah, A.H.J.
 Biochem. Biophys. Res. Commun. 251, 252-255, 1998
 A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidops
 A:Reference number: JE0369; MUID:99009256
 A:Accession: JE0369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-465 <MOL>
 A:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 6.3% Score 156; DB 2; Length 465;
 Best Local Similarity 22.2%; Pred. No. 0.00012;
 Matches 100; Conservative 65; Mismatches 200; Indels 86; Gaps 20;

QY 35 DGGYQCPEISHLMGOVSPFFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTSSSKK 94
 Db 16 DGGFDVRHHLSTV-TRYSTSKDVTQNLIEGNSVPSECTPHLNLVARHGTRSPKTKLRE 74
 QY 95 VSAL--TEATOKNATAPK-----GKAFIKTVNTLGGADLTTPFGEOQMVNSGI 141
 Db 75 MESLAGREFELVRDAEARKLSKDPKIPWLGOWK--SPWEGKVKGELLROGEDELYQLGI 132
 QY 142 KFYRYKALARK---IVPFVRASGSDRVIASAEKFTIEGFOSAK-LADPGANPHQASPIV 196
 Db 133 RYRERFSLFEEDYHPDVYVTRATQIPRASASAVAFGMGLSEKGNLGPGRNRAFA---- 188
 QY 197 NVIIPEGAGYNTLHGLCTAFEESELDGDDVEANFTAVFAPPIRARLEAHLPGVNLTD 256
 Db 189 --VTSENRASTDKLRFEC-----QNY-----KSYRKAKPEAVDKLKEP 226
 QY 257 VVNLMDMC---PFDTVARTSDATQSPFCDLTFHDEWTDYDYLQSLGKYGYGAGNPLGP 313
 Db 227 VLNKITASVAKRYDKLFTKQDISLWFLCKQVALLW-----TDDLEVFLLKGYGNSLNY 281
 QY 314 AGQGVFNVELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFFA-L 372
 Db 282 KMGVPLLEDVL---HSMEEAIKAREKLP--PGSYE-KARL--RFAHAETIVPPFSCLL 332
 QY 373 GLYNGTKPLSTTSVESIE-----ETDGYAASWTVPFAARAYVEMMOCEAEK 420
 Db 333 GLF-----LDGSEFEKIQEKEPLELPQPPKTRDFRGSTMAPFGGNLILVLYSCPAESSP 387
 QY 421 --LVRLVNDVVRVPLHGGCGVDKLGKRDFF 449
 Db 388 KYFVQVHLNHEPIAVPGC--DGKDFCPLDEF 416

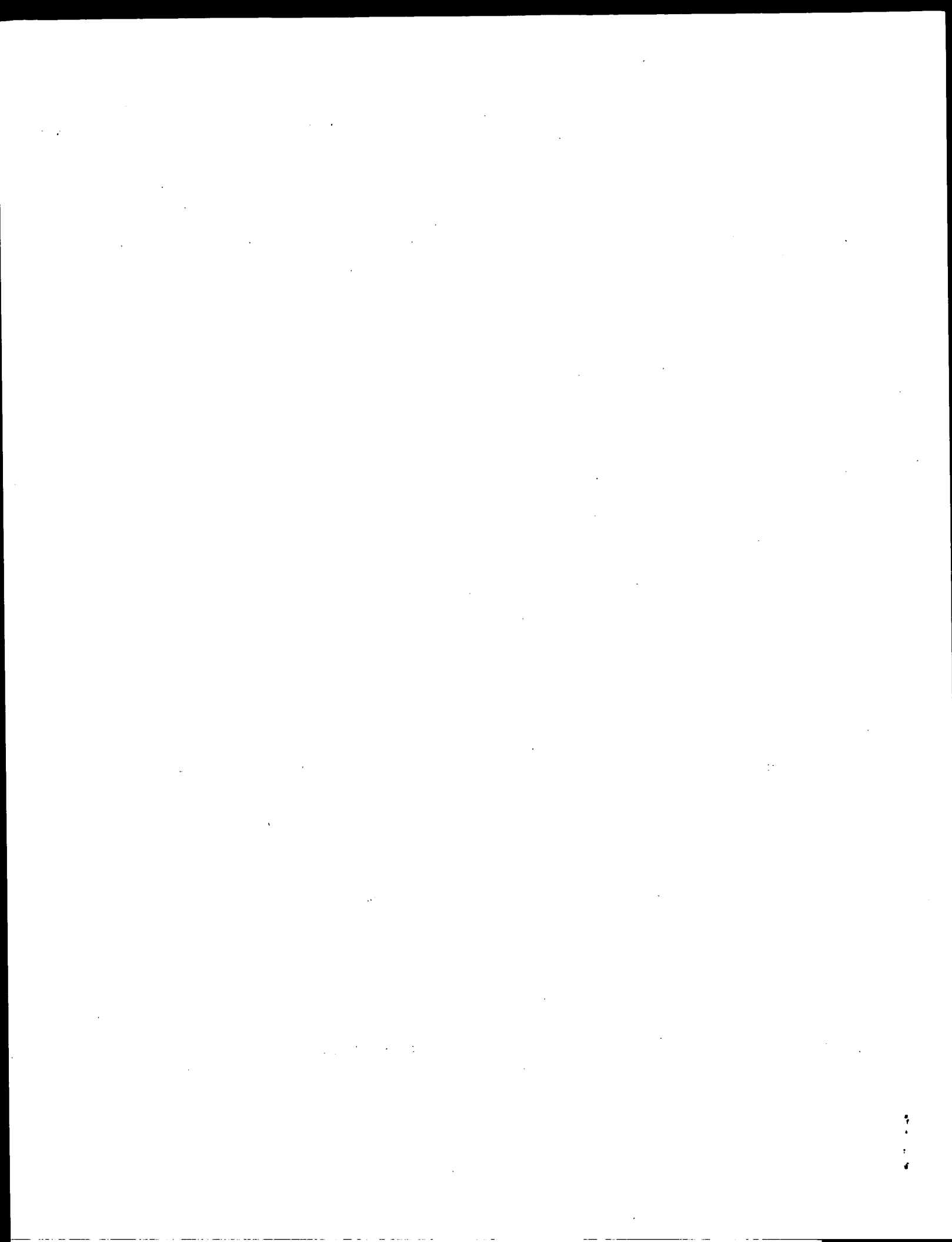
Search completed: October 26, 2001, 15:10:27
 Job time: 6333 sec

QY 51 YSPFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTSS-----SKSKYSALIE----- 100
 Db 43 HEPYF-----NGPTTTSFESCAIKQVHLLQHRGSRNPTGDDTATDVSSAQYIDIFQKLL 97
 QY 101 --AIQKNATAFKGYAFKTYNTL---GADLTTPFGEOQMVNSGIKRYRYKALARKIV 155
 Db 98 NGSIPVNFSEYPENPLFYFKHMTFPVKAENADQLSSSGRIELFDLGRQVFERIYELFDYDV 157
 QY 156 PFVRASGSDRVIASAEKFTIEGFOSAKLADPGANPHQASPVINVIPE--GAGYNTLHDG 213
 Db 158 YDINTAQERVVDSAEWFSYGM-----FGDDMQNKTNF-----IVLPEDDSAGANSLSMY 208
 QY 214 LCTAFEESELGDD-VEANFTA---VFAPPPIRARLEAHL-PGVNLTDDEVNLMDCPFD 268
 Db 209 SCPVYEDNNIDENTTAAHTSRNFKLPIANRLNKYSDSGYNLTVSDVRSLYYICVYEI 268
 QY 269 VARTSDATQSPFCDLTFHDEWTDYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLT 328
 Db 269 ALRDN-----SDFCSLFTPTSEFLNFEYDSDLQYAIWGGPASEWASTLGCAYVNNLANL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFALGLYNGTKPLSTTSVES 388
 Db 323 -----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDITPEHPLPTDK 367
 QY 389 IEETDGYAASWTVPFAARAYVEMMOCEAEKPLRVLRVNDVVRVPLHGGCGVDKLG 442
 Db 368 NIYYSOKTSFVPFAGNLITELFFC-SDSKYYVRHLVNOQVYPLDLCGYPGSG 420

RESULT 14
 T39929
 thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39929
 R: Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.
 submitted to the EMBL data Library, May 1998
 A:Reference number: Z21857
 A:Accession: T39929
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-463 <LYN>
 A:Cross-references: EMBL:AL023286; PIDN:CAAL8963.1; GSPDB:GN00067; SPDB:SPBC21H7.03c
 A:Experimental source: strain 972h; cosmid c21H7
 C:Genetics:
 A:Gene: SPDB:SPBC21H7.03c
 A:Map position: 2
 C:Superfamily: yeast acid phosphatase

Query Match 12.2% Score 301; DB 2; Length 463;
 Best Local Similarity 24.3%; Pred. No. 2.7e-16;
 Matches 106; Conservative 71; Mismatches 198; Indels 62; Gaps 13;

QY 51 YSPFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTS--SKSKYSA-----L 98
 Db 43 HEPYFDGLDSA-----FPTEICIQVHLLQHRGSRNPTGDDTATDVSSQYLNNEFQKLL 97
 QY 99 IEATOKNATAFKGYAFKTYNTL---GADLTTPFGEOQMVNSGIKRYRYKALARKIV 155
 Db 98 NGSIPVNFSEYPENPLCFIKTQVTPVIDAENADQLSSSGRIELFDLGRQVRYKLFDSYV 157
 QY 156 PFVRASGSDRVIASAEKFTIEGFOSAKLADPGANPHQASPVINVIPEG--AGYNTLHDG 213
 Db 158 YDINTAQERVVESAKWTFYGLFGKWE-----KTNFILLSGKAAGANSLSMYN 208
 QY 214 LCTAFEESEL-----GDDVEANFTAVFAPPPIRARLEAHL-PGVNLTDDEVNLMDCPFD 268
 Db 209 ACPVFKDNFNHFNKATDAHAHVRNIFIEPIVNLAKYFSDSSYKLTINDVRSFLYICEYEI 268
 QY 269 VARTSDATQSPFCDLTFHDEWTDYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLT 328



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:34 ; Search time 29.88 Seconds
(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-26

Perfect score: 2470
Sequence: 1 MGVEVLLSIATLFGTSGT.....DFVEGLSFARSGGNWEECA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1866	75.5	467	1	PHYA_ASPAW
2	1862	75.4	467	1	PHYA_ASPNG
3	1813.5	73.4	463	1	PHYB_EMENI
4	403	16.3	467	1	PPA5_YEAST
5	399	16.2	467	1	PPAB_YEAST
6	399	16.2	467	1	PPAC_YEAST
7	394	16.0	467	1	PPA3_YEAST
8	370	15.0	468	1	PPAD_YEAST
9	352.5	14.3	479	1	PHYB_ASPNG
10	350	14.2	468	1	PPAL_PICPA
11	346.5	14.0	479	1	PHYB_ASPAW
12	340.5	13.8	469	1	PPA5_KLULA
13	326	13.2	453	1	PPA1_SCHPO
14	301	12.2	463	1	PPA2_SCHPO
15	134	5.4	755	1	PPAX_CAEEL
16	122.5	5.0	423	1	PPAL_RAT
17	117.5	4.8	421	1	PPAL_MOUSE
18	112.5	4.6	423	1	PPAL_HUMAN
19	108.5	4.4	381	1	PPAP_RAT
20	107.5	4.4	602	1	PERX_ERWCH
21	104.5	4.2	413	1	AGP_ECOLI
22	100.5	4.1	693	1	LYS4_YEAST
23	98	4.0	386	1	PPAP_HUMAN
24	98	4.0	844	1	YD25_SCHPO
25	97	3.9	1350	1	VG72_HSVII
26	96	3.9	721	1	NICA_CAEEL
27	95.5	3.9	433	1	DHOM_SVNY3
28	95.5	3.9	609	1	COPA_PSESM
29	95	3.8	938	1	TRGL_ECOLI
30	95	3.8	3587	1	TYCB_BACRI
31	94	3.8	411	1	PPAY_CAEEL
32	94	3.8	467	1	ATPX_BACFI
33	94	3.8	1527	1	MRP3_HUMAN

34 93.5 3.8 396 1 ARSL_AQUAE
35 93.5 3.8 646 1 IRG1_MOUSE
36 93.5 3.8 2485 1 PTND_HUMAN
37 91.5 3.7 526 1 CLOS_CLOHT
38 91.5 3.7 586 1 VP40_ILTVT
39 91 3.7 413 1 PPAB_CAEEL
40 90 3.6 366 1 ODPB_YEAST
41 90 3.6 428 1 GSA_STAAU
42 90 3.6 626 1 BGAL_LEULA
43 89.5 3.6 845 1 CC47_YEAST
44 89 3.6 417 1 AGP_PRORE
45 89 3.6 1191 1 DNBI_MCMVS

ALIGNMENTS

RESULT 1
ID PHYA_ASPAW STANDARD; PRT: 467 AA.
AC P34753;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA OR PHY.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALK0243;
RX MEDLINE=94040796; PubMed=8224894;
RA Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
RA Miettinen-Oinonen A., Nevalainen H., Rambossek J.A.;
RT "The cloning and sequencing of the genes encoding phytase (phy) and
RT pH 2.5 optimum acid phosphatase (aph) from Aspergillus niger var.
RT awamori.";
RL Gene 133:55-62(1993).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L02421; AAA16898.1;
CC PIR: JN0889; JN0889.
CC HSP: P34752; LIHP.
CC InterPro: IPR000560;
CC Pfam: PF00328; acid_phosphat; 1.
CC PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC Hydrolase; Glycoprotein; Signal.
CC SIGNAL 1 23 POTENTIAL.
CC CHAIN 24 467 3-PHYTASE A.
CC ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
CC ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
CC DISULFID 31 40 BY SIMILARITY.
CC DISULFID 71 41 BY SIMILARITY.
CC DISULFID 215 465 BY SIMILARITY.
CC DISULFID 264 282 BY SIMILARITY.
CC DISULFID 436 444 BY SIMILARITY.

```

FT CARBOHYD 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 467 AA; 51075 MW; 118E28A5D7EC661 CRC64;

Query Match 75.5%; Score 1866; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 3.5e-143;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVSFVLLSLATLFGSTGSLALPGRNHSCHSDTVGGVQCFFPEISHLWGQYSPFFSLADE 60
Db 1 MGVSFVLLSLATLFGSTGSLALPGRNHSCHSDTVGGVQCFFPEISHLWGQYSPFFSLANE 60
QY 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALTEAIOKNATAPKGYAFLKTYN 120
Db 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALTEAIOKNATAPKGYAFLKTYN 120
QY 121 YTLGADLTPFGQOMVNSGKIFRYKALARKIVPFRASGSDRVASAEKFIQFQSA 180
Db 121 YTLGADLTPFGQOMVNSGKIFRYKALARKIVPFRASGSDRVASAEKFIQFQSA 180
QY 121 YSLGADLTPFGQOMVNSGKIFRYKALARKIVPFRASGSDRVASAEKFIQFQSA 180
Db 121 YSLGADLTPFGQOMVNSGKIFRYKALARKIVPFRASGSDRVASAEKFIQFQSA 180
QY 181 KLADPGANPHQASPNVNIPEGAGVNTLDHGLCTAPFESELGDDVEANFTAVAPPTIR 240
Db 181 KLADPGANPHQASPNVNIPEGAGVNTLDHGLCTAPFESELGDDVEANFTAVAPPTIR 240
QY 181 KLADPGANPHQASPNVNIPEGAGVNTLDHGLCTAPFESELGDDVEANFTAVAPPTIR 240
Db 181 KLADPGANPHQASPNVNIPEGAGVNTLDHGLCTAPFESELGDDVEANFTAVAPPTIR 240
QY 241 ARLEALHPGVNLTDEVDVNLMDMCPEDTIVARTSDATQSLPFCDFLTHDEWIOYDLQSLG 300
Db 241 ARLEALHPGVNLTDEVDVNLMDMCPEDTIVARTSDATQSLPFCDFLTHDEWIOYDLQSLG 300
QY 241 QRLNDLSGVTLDTEVTLMDMCPEDTIVARTSDATQSLPFCDFLTHDEWIOYDLQSLG 300
Db 241 QRLNDLSGVTLDTEVTLMDMCPEDTIVARTSDATQSLPFCDFLTHDEWIOYDLQSLG 300
QY 301 KYGYGAGNPLGPAQGVNFVNLIELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
Db 301 KYGYGAGNPLGPAQGVNFVNLIELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
QY 301 KYGYGAGNPLGPAQGVNFVNLIELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
Db 301 KYGYGAGNPLGPAQGVNFVNLIELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
QY 361 HDNTMVSIFFALGLNGTKPLSTSVESIEDTGYAASWTVPFAARAYVEMMOCEAEKEP 420
Db 361 HDNTMVSIFFALGLNGTKPLSTSVESIEDTGYAASWTVPFAARAYVEMMOCEAEKEP 420
QY 421 LVRVLVNDVRVPLHGGCVKGLRCKRDKDFVEGLSFRASGSGNWEECEFA 467
Db 421 LVRVLVNDVRVPLHGGCVKGLRCKRDKDFVEGLSFRASGSGNWEECEFA 467

```

```

RESULT 2
PHYA_ASPPG STANDARD; PRT; 467 AA.
AC P34752;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCR-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93252284; Pubmed=8387447;
RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
RA Gouka R.J., Suykerbuyk M.E.G., Luiten R.G.M., van Paridon P.A.,
RA Selten G.C.M., Veenstra A.E., van Gorcom R.F.M.,
RA van den Hondel C.A.M.J.J.;

```

```

RT "cloning, characterization and overexpression of the phytase-encoding
RT gene (phyA) of Aspergillus niger.";
RN Gene 127:87-94 (1993).
RN [2]
RN SEQUENCE FROM N.A.
RA Mullanev E.J.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 24-464.
RP STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93249451; Pubmed=8387289;
RA Ullah A.H.J., Dischinger H.C. Jr.;
RT "Aspergillus ficuum phytase: complete primary structure elucidation
RT by chemical sequencing.";
RN Biochem. Biophys. Res. Commun. 192:747-753 (1993).
RN [4]
RN SEQUENCE OF 71-93.
RP STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=91298982; Pubmed=1648914;
RA Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
RT "Cyclohexanedione modification of arginine at the active site of
RT Aspergillus ficuum phytase.";
RN Biochem. Biophys. Res. Commun. 178:45-53 (1991).
RN [5]
RN CHARACTERIZATION, AND PARTIAL SEQUENCE.
RP STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=89160685; Pubmed=2852807;
RA Ullah A.H.J.;
RT "Aspergillus ficuum phytase: partial primary structure, substrate
RT selectivity, and kinetic characterization.";
RN Prep. Biochem. 18:459-471 (1988).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RP STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=97307250; Pubmed=9164457;
RA Kostrewa D., Gruninger-Leitch F., D'Arcy A., Broger C., Mitchell D.,
RA van Loon A.P.;
RT "Crystal structure of phytase from Aspergillus ficuum at 2.5-A
RT resolution.";
RN Nat. Struct. Biol. 4:185-190 (1997).
RN [7]
RN FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
RN FROM PHYTASE.
RN CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
RN D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
RN BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN
RN FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
RN USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER
RN THE NAME PHYTASE NOVO.
RN [8]
RN SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
RN [9]
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN [10]
RN EMBL: Z16414; CAA78904.1; .
RN EMBL: M94550; AAA32705.1; .
RN PIR: JN0482; JN0482.
RN PIR: JN0023; JN0023.
RN PIR: JN0656; JN0656.
RN PDB: 1HP; 18-MAR-98.
RN InterPro: IPR000560; .
RN Pfam: PF00328; acid.phosphat. 1.
RN PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
RN PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
RN Hydrolase; Glycoprotein; Signal; 3D-structure.
RN SIGNAL 1 23
RN CHAIN 24 467 3-PHYTASE A.
RN ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
RN ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).

```

```

FT DISULFID 31 40
FT DISULFID 71 414
FT DISULFID 215 465
FT DISULFID 264 282
FT DISULFID 436 444
FT CARBOHYD 27 27
FT CARBOHYD 59 59
FT CARBOHYD 105 105
FT CARBOHYD 120 120
FT CARBOHYD 207 207
FT CARBOHYD 230 230
FT CARBOHYD 339 339
FT CARBOHYD 352 352
FT CARBOHYD 376 376
FT CARBOHYD 388 388
SQ SEQUENCE 467 AA; 51086 MW; 88FE9F3584341D6D CRC64;

Query Match 75.4%; Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 7.3e-143;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGPRGNHSCDVTVDGGYQCPEISHLMGQYSPFFSLADE 60
DB 1 MGVSALLPLVLLSGVTSGLAVPASRNSSCDTVDGGYQCPEISHLMGQYAPFFSLANE 60
QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGKYAFKLTYN 120
DB 61 SVISPEVAGCRVTFQAQVLSRHGARYPTDSKSKYSALIEAIQKNATAFKGKYAFKLTYN 120
QY 121 YTLGADLLTPGEOQVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGFOSA 180
DB 121 YSLGADLLTPGEOELVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGFOSA 180
QY 181 KLADPGANPHQASPVNIIPEGAGYNTLDHGLCTAFPESELGDDVEANTAVFAPP 240
DB 181 KLDPRAQPGSSPKIDVIVSEASSNNITDPGCTVFEDSELADTVEANTAVFAPP 240
QY 241 ARLEAHLPGVNLTDDEVVNLMDCPFDTVARTSDATQSLSPCDLFTHDEWIQYDYLQSLG 300
DB 241 ORLENDLSGVTLTDEVTYLMDCMSFDTISTSTVDTKLSLPCDLFTHDEWIQYDYLQSLG 300
QY 301 KYGAGNPLGPAQGVGVNELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYCHGAGNPLGTPQGVYANELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
QY 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYASAWTPVFAARAYVEMMQCEAKEP 420
DB 361 HDNGIISILFALGYNKPLSTTTVENITQTDGFSANWTPVFAARAYVEMMQCEAKEP 420
QY 421 LVRVLVNDVRVPLHGCVKLGRCRDRDFVEGLSFARSGGNWCECFA 467
DB 421 LVRVLVNDVRVPLHGCPVDALGRCTRDSFVRGLSFARSGGDNWCECFA 467

```

RESULT 3

```

PHYB_EMENI
ID PHYB_EMENI STANDARD; PRT; 463 AA.
AC O00093;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE
DE 3-PHOSPHOHYDROLASE B)
DE PHOSPHOHYDROLASE B)
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98007872; PubMed=9349716;
RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,

```

```

RA van Loon A.P.;
RT "Cloning of the phytases from Emericella nidulans and the
RL thermophilic fungus Talaromyces thermophilus.";
RL Biochim. Biophys. Acta 1353:217-223(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U59803; AAB96871.1; -.
DR InterPro: IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 463
FT ACT_SITE 80 80
FT ACT_SITE 81 81
FT ACT_SITE 357 357
FT CARBOHYD 26 26
FT CARBOHYD 41 41
FT CARBOHYD 103 103
FT CARBOHYD 118 118
FT CARBOHYD 203 203
FT CARBOHYD 226 226
FT CARBOHYD 331 331
FT CARBOHYD 335 335
FT CARBOHYD 372 372
SQ SEQUENCE 463 AA; 51786 MW; ECC5827DIEIC82A2 CRC64;

```

0;

```

Query Match 73.4%; Score 1813.5; DB 1; Length 463;
Best Local Similarity 72.7%; Pred. No. 5.9e-139;
Matches 339; Conservative 48; Mismatches 74; Indels 5; Gaps 3;

```

```

QY 1 MGVFVLLSIATLFGSTGALGPRGNHSCDVTVDGGYQCPEISHLMGQYSPFFSLADE 60
DB 1 MAFTVALSLYLLSRVSAQA--PVQNHSCNTADGGYQCPEPNVSHVMGQYSPVFSIQE 58
QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGKYAFKLTYN 120
DB 59 SAISEDVPHGCEVTFVQVLSRHGARYPTESKSKAYSGLIEAIQKNATFSWQYAFLESYN 118
QY 121 YTLGADLLTPGEOQVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGFOSA 180
DB 119 YTLGADLLTPFGENQWVDGSAKEFYRYKALARKIVPFIRASGSDRVIASAEKFTIEGFOSA 178
QY 181 KLADPGANPHQASPVNIIPEGAGYNTLDHGLCTAFPESELGDDVEANTAVFAPP 240
DB 179 QLDHGHGSK--RATPVNVIIPEDGFNNTLDHSCVSENDERADEATEANFTALMGPP 236
QY 241 ARLEAHLPGVNLTDDEVVNLMDCPFDTVARTSDATQSLSPCDLFTHDEWIQYDYLQSLG 300
DB 237 KRLNDLPGLKLNENVIYLMDCSFDTMARTAGHTSELSFCALFTKEKWLQYDYLQSLG 296
QY 301 KYGAGNPLGPAQGVGVNELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 297 KYGAGSPLGPAQGVGVNELLARLTHSPVDHTSTNHTLDSNPATFPLDRKLYADES 356
QY 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYASAWTPVFAARAYVEMMQCEAKEP 420
DB 357 HDNSMISIFFAMGLYNGTQPLSMDSVESIQEMDGYAASWTPVFGARAYFELMQCE-KKEP 415

```

QY 421 LVRVLNDRVPLHGGVDKLGCRKRDFFVEGLSFARSGGNWEECF 466
 Db 416 LVRVLNDRVPLHGGVDKLGCRKRDFFVEGLSFARSGGNWEECF 461

RESULT 4
 PPA5_YEAST STANDARD; PRT; 467 AA.
 AC P00635;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
 GN PHO5 OR YBR093C OR YBR0814.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 18-26.
 RX MEDLINE=83168913; PubMed=6300772;
 RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
 RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
 of repressible acid phosphatase contains a signal peptide.";
 RL Nucleic Acids Res. 11:1657-1672(1983).
 RN [2]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 18-45.
 RX MEDLINE=85037940; PubMed=6093051;
 RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RT "Structural analysis of the two tandemly repeated acid phosphatase
 genes in yeast.";
 RL Nucleic Acids Res. 12:7721-7739(1984).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R.; Ehle S., Vetter I., Feldmann H.;
 RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 RN [4]
 RP SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE=87064474; PubMed=3537710;
 RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
 RA Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
 RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
 cluster within the acid phosphatase multigene family of Saccharomyces
 cerevisiae.";
 RL Mol. Cell. Biol. 6:1855-1865(1986).
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
 CC -1- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; V01320; CAA24630.1;
 CC EMBL; X01079; CAA25555.1;
 CC EMBL; X78993; CAA55598.1;
 CC EMBL; Z35962; CAA85046.1;
 CC EMBL; X01080; CAA25556.1;
 CC PIR; S05795; PABYC.
 CC PIR; B25241; B25241.
 CC HSP; P34752; LIHP.
 CC SGD; S0000297; PHO5.

DR InterPro: IPR000560;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 18 467 REPRESSIBLE ACID PHOSPHATASE.
 FT ACT_SITE 75 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT CARBOHYD 97 97 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 36 36 D -> Y (IN REF. 1).
 FT CONFLICT 130 130 H -> G (IN REF. 1).
 FT CONFLICT 294 294 H -> Q (IN REF. 1).
 FT CONFLICT 446 446 S -> V (IN REF. 1).
 FT CONFLICT 462 463 AS -> DT (IN REF. 1).
 FT CONFLICT 466 466 R -> K (IN REF. 1).
 SQ SEQUENCE 467 AA; 52858 MW; DC3G9504BC2D3D0C CRC64;

Query Match 16.3%; Score 403; DB 1; Length 467;

Best Local Similarity 26.2%; Pred. No. 4.8e-25;
 Matches 126; Conservative 61; Mismatches 216; Indels 78; Gaps 15;

QY 6 VLLSTATLFGSTALPGNGNSHSDTVGGYQCFPEISHLWGQYSPFFSLADESAISP 65
 Db 5 VVYSLAASLANAGTI--PLGKLADVDKIGTKQDIFPL---GGAGPYVSPFGDYGISR 58
 QY 66 DVPKGRVTVQVLSRHGARYPTSSKKYSALIEATOKNATAPKGYAFLK-TYNTLGL 124
 Db 59 DLPECEMKQLQMVGRHGERYPYTVSLAKTIKSTWYKLSNTYTRQFNGLSFLNDDYEFFIR 118
 QY 125 ADD-----LTPF-GEQMVNSGKIFRYRKALARKIVPF-VRASGSDRVIA 168
 Db 119 DDDLEMETTFANSDVDVLPYTGEMNAKRHARDELAQYGYAVENQTSFAVTSNKRCHD 178
 QY 169 SAEKPIEGFQSAKLADPGANPHQASPVINII-----PEGAGYNNTLDHGLCTAFEESEL 223
 Db 179 TQAYFIDG-----LGDQ-----FNITLQTVSEASAGANTLSACNSCPAW-DYDA 222
 QY 224 GDDVEANFVAFPIRARELAHLPGVNLTDVVDVMDMCPFTVARTSDATOLSPFCD 283
 Db 223 NDDIVNEYDTTTLDDIAKRLNKENKGLNLTSTDASTLFSWCAFEV-----NAKGYSDVC 277
 QY 284 LFTDEWLOQYDLOSIGKYCYGAGNPLGPAQGVGFVNELIARLTHSPVOOHTSTNHTLD 343
 Db 278 IPTKDELVHSYQDLHTYYHGGPGYDIKSVGNLNFASVKLLKQSEIQDQ----- 329
 QY 344 SNAPTFPLNATLYADFSHDNTWSVIFFFALGLYNGTKPLSTTSVESIETDGYAASWTVPF 403
 Db 330 -----KWLVSFTHTDILNFLTLAGIIDDKNLNTAEYVPMGMT--FHRSWYVPQ 377
 QY 404 AARAVEMMQCEAEKPELVRLVNDVRVPLHGGVDKLGCRKRDFF-----VEGLSFA 456
 Db 378 GARVYTERFQC--SNDTYRVYVINDAVVPIETCTSGPGFSCSEINDFYDYAEKRVAGTDFL 435
 QY 457 R 457
 Db 436 K 436

RESULT 5
 PPAB_YEAST

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

01-MAR-1992 (Rel. 21, Created)
01-OCT-1994 (Rel. 30, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
PHO3 OR YBR092C OR YBR0813.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
MEDLINE=85037940; PubMed=6093051;
Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
"Structural analysis of the two tandemly repeated acid phosphatase
genes in yeast";
Nucleic Acids Res. 12:7721-7739(1984).
[2]
SEQUENCE FROM N.A.
STRAIN=S288C;
MEDLINE=95208357; PubMed=7900426;
Mannhaupt G., Stucka R., Ehmlé S., Vetter I., Feldmann H.;
"Analysis of a 70 kb region on the right arm of yeast chromosome II.";
Yeast 10:1363-1381(1994).
-!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
ALCOHOL + ORTHOPHOSPHATE.
-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
GN PHO3 OR YBR092C OR YBR0813.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85037940; PubMed=6093051;
RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
RT "Structural analysis of the two tandemly repeated acid phosphatase
genes in yeast";
RL Nucleic Acids Res. 12:7721-7739(1984).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95208357; PubMed=7900426;
RA Mannhaupt G., Stucka R., Ehmlé S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
RL Yeast 10:1363-1381(1994).
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

EMBL; X01080; CAA25557.1; -
DR EMBL; X78993; CAA55597.1; -
DR EMBL; Z35961; CAA85045.1; -
DR PIR; S44674; PABYCC.
DR HSP; P34752; 1IHP.
DR SGD; S0000296; PHO3.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat_1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17
FT CHAIN 18 467
FT ACT_SITE 75 75
FT ACT_SITE 337 337
FT CARBOHYD 97 97
FT CARBOHYD 103 103
FT CARBOHYD 162 162
FT CARBOHYD 192 192
FT CARBOHYD 250 250
FT CARBOHYD 315 315
FT CARBOHYD 356 356
FT CARBOHYD 390 390
FT CARBOHYD 439 439
FT CARBOHYD 445 445
FT CARBOHYD 456 456
FT CARBOHYD 461 461
FT CONFLICT 219 221
FT SEQUENCE 467 AA; 52776 MW; 05FBB80DEB41B0FF CRC64;
Query Match 16.2%; Score 399; DB 1; Length 467;
Best Local Similarity 26.1%; Pred. No. 1e-24;
Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;
QY 24 PRGNSCHDVTGGYQCFPEISHLWGOYSPFFSLADESAISPDVPGCRVTFVQVLSRHG 83
DB 21 PLGKLSIDDKTGTEIPFPL----GGSGPYSPGPGYIGISRDLPESCEMKQVMYGRHG 76
QY 84 ARYPTSSKSKYSALIEAIOKNATAFRGKYAFK-----TYNYTL-----GADDL 128
DB 77 ERYPTVSKAKSMITWYKLSNYTGFSGALSFLNDDYEFFTRDTKNEMETTLANSVNL 136
QY 129 TPF-GEQMVNSGKIFRYRYKALARKIVPF-VRASGSDRVTSASAEKIEGFQSKADPG 186
DB 137 NPYTGEMNAKRHRDFLAQYGYMVENOTSFAVFTSNRCHDTAQYFDG-----LGDK- 190
QY 187 ANPHQASPVNII-----PEGAGVNNLTDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241
DB 191 -----FNISLOTISEASAGANTLSAHSFPAWDD-DVNDLKKYDTKYLSTGIAK 240
QY 242 RLEAHLPGVNLTDDEVNMDMCPDVTARTSDATQLSFPDCLTHDEWIOYDYLQSLGK 301
DB 241 RLNKNKGLNLTSSDNTFFAWCAYEINAR-----GYSDICNIPTKDLVRFSTYGGDLET 295
QY 302 YGYGAGNPLGPAQGVGFVNELIARLTHSPQDHTSTNHTLDSNPATPEPLNATLYADF 361
DB 296 YYTGPGYDVVRSGANLNFASVKKLKESEVQDQ-----KWLSEFTH 337
QY 362 DNTMYSIFFALGLYNGTKPLSTSVESIEEDTGYAASWTVPFAARAYVEMMQCEKEPL 421
DB 338 DTDILNLTGIIIDQNNLTAHEHPFMENT--PHRSWYVFGQARVYTEKFC--SNDTY 393
QY 422 VRVLVNDRVVPLHGGVDKLGCRKDDF-----VEGLSFAR 457
DB 394 VRYVINDAVVPIETCTGPGFSCFNDYFGYAEKRVAGTDFLK 436
RESULT 7
PP33_YEAST
ID PP33_YEAST
AC P24031;
PRT; 467 AA.

Query Match 16.0%; Score 394; DB 1; Length 467;
Best Local Similarity 27.0%; Pred. No. 2.5e-24;
Matches 117; Conservative 62; Mismatches 191; Indels 64; Gaps 13;
QY 49 QOYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATA 108
DB 42 GGAGPYSPFGDYGISRDLPESCEMKQVMYGRHYPTYSKGTATMTWYKLSNYTRQ 101

QY 109 FKGYAFLEK-TYNTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFUNDIEFFIRDDDDLEMETTFANSNDVNLNPTGEMDAKRHAREFLAQGYWFE 161
 QY 153 KIVPF-VRASGSDRVASAEKFIIEGFSQAKLADPGANPHQASPVNLIIPGAGYNNITLD 211
 Db 162 NOTSFPIFAASSERVDHTAQYFIDGL-----GDQFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFEESELGDDVEANTAVFAPPFIRARLEAHLPGVNLTDVVDVMDMCPFDITVA 270
 Db 211 AGNACPDWDE-DANDDLKDYDTYLDLIDIAKRLNKENKGLNLTSKDANTLFAWCAYELNA 269
 QY 271 RTSDATOLSPFCDLFTDHEWIOYDLOSGLKGYGAGNPLGPAQGVGVFNELIARLTHS 330
 Db 270 R-----GYSVDCTIDTEDELVRYSYGDLVSFYODGPGYDMIRSVGANLNFATLKLLKQS 324
 QY 331 PVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFGALGLYNGTKPLSTSVESIE 390
 Db 325 ETQD-----LKVLSFTHDIDLNLTLTAGIIDDKNNLTAEYVFFMG 366
 QY 391 ETQGYAASVTPFAARAYVEMMOCEAEKEPLVRVNLVNDVVPVPLHGGCVGDKGRCKRDDF- 449
 Db 367 NT-FHKSWVPOGARVYTEKFC--SNDTYVRVINDAVVPIETCSTGPGFSCINDFY 422
 QY 450 -----VEGLSFAR 457
 Db 423 DYAEKRVAGTDFLK 436

RESULT 8

PPAD_YEAST STANDARD; PRT: 468 AA.
 ID PPAD_YEAST
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROBABLE ACID PHOSPHATASE YDL024C PRECURSOR (EC 3.1.3.2).
 GN YDL024C OR D2615.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Vissers S., Urrestazu L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; Z48432; CAAB8335.1; -
 DR EMBL; Z74072; CAA98583.1; -
 DR SGD; S0002182; YDL024C.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid.phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 20 BY SIMILARITY.
 FT CHAIN 21 468 PROBABLE ACID PHOSPHATASE YDL024C.
 FT ACT_SITE 76 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 338 338 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 163 163 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDF162C CRC64;
 Query Match 15.0%; Score 370; DB 1; Length 468;
 Best Local Similarity 25.1%; Pred. No. 2.2e-22;
 Matches 116; Conservative 62; Mismatches 174; Indels 110; Gaps 14;
 QY 49 GOVSPFFSLADESAISPDVPKGRVTEVQVLSRHRGARYPTSSKSKYSALTEAIQKNATA 108
 Db 43 GGSAPYFSPFANYGIPDIPEGCRLTQVMIGRHGERYPTRSEAKDIFEVYKISNTYTK 102
 QY 109 FKGYAFLEK-TYNTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 103 YEGSLFLNNGYEFFIPDESLEMETTLQNSIDVNLNPTGEMDAKRHAREFLAQYKLM 162
 QY 153 KIVPF-VRASGSDRVASAEKFIIEGFSQAKLADPGANPHQASPVNLIIPGAGYNNITLD 211
 Db 163 NCTNPFITNSKRIVDTAQYFAEL-----GDGFNLSLQ 197
 QY 212 HGLCTAFEESELG-----DDVEANTAVFAPPFIRARLEAHLPGVNL 253
 Db 198 -----TLSENSSSGANTLAAKSSCPNNSNANDILMSYSDYLENISDRLDENKGLNLS 253
 QY 254 DEDVNLMDMCPEDTAVRTSDATOLSPFCDLFTDHEWIOYDLOSGLKGYGAGNPLGP 313
 Db 254 RKDAALFNSWCAPEL-----NAKYSNIDCFSAELIHSYETDITSTFYONGPGYKLIK 308
 QY 314 AQGVGVN---ELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADESHDNTMVSIF 370
 Db 309 SGANLNFATVKLIQSAH-----LDQKWLSTHTDIDLNYLT 347
 QY 371 ALGLYNGTKPLSTSVESIEETDGYAASVTPFAARAYVEMMOCEAEKEPLVRVNLVND 430
 Db 348 TAGLIDTNRNLTNHVPFRDHS--YHRSWVIOGARVYTEKFC--SNDSYVRVYVND 403
 QY 431 VPLHGGCVGDKGRCKRDDFVE-----GLSFARSGNWEEC 465
 Db 404 VPIESCGSGPGFSCERGTEFYAKDRLCGVSF-----YEDC 439

RESULT 9

PHVB_ASPNG STANDARD; PRT: 479 AA.
 ID PHVB_ASPNG
 AC P34754;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE B).
 GN PHVB.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.
 RX MEDLINE-93371452; PubMed=7916610;
 RA Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,
 RA Ullah A.H.J.;
 RT "Identification and cloning of a second phytase gene (phyB) from
 FT Aspergillus niger (ficcum).";
 RL Biochem. Biophys. Res. Commun. 195:53-57(1993).

[illegible]

Query Match	14.2%	Score	350;	DB	1;	Length	468;
Best Local Similarity	25.4%;	Pred. No.	9.1e-21;				
Matches	125;	Conservative	79;	Mismatches	199;	Indels	90;
Gaps	21;						
Qy	6	VLLSIATL--	-FGSTG	TALGPRGN	SHSCD	TDG	GYOCQFPEISHLWQYSPF--FSLADE
Db	10	IILALATLQSVAFELQHV	LGVDNR	DPY	PQRT--	DDQY	NILRHLLGGL- GPYICYNQWGIAAE
Qy	61	SAISPDVPKGRVTFVQV	LSRHGARY	TSSKSKYS	ALIE--	AIQ	NATAFKGKYAFIKTY
Db	68	SEI-----	ESCTID	OAHLMLRHGERY	FTSNV	GKQLEALY	QKLLDADVEVTPGLPSFFQDY
Qy	120	NYTLG----	ADDLTPF	GQQMVNSG	IKF----	YRRYKAL-----	AKIVPFVRASSDR
Db	123	DYFSDAAWYEOETT	KGY	SGLNTAF	QFTGLRERY	DHLLINTEG	KKL--SVWAGSOER
Qy	166	VIASKEFI	EGFOSAKL	DPGAN	PHQASPV	TNVIPE---	GAGYNNTLDHGLCT-----A
Db	181	VDTATYFAQFMKSN	NYTD-----	MVEV	VALEEK	SGQLNSLTARIS	CPNYSNHI
Qy	218	FEESLGGDV---	EANFTAV	APPIR	ALEAHL	PGVNLTD	DVVNLMDCPFDTVARTSD
Db	231	YKDGDFPN	DIAREAD-----	RLNT	SPG	FNITAD	DIPTIYALCYGFLNVRCG-

14 28: score 350: DB 1: Length 468;

14 28: score 350: DB 1: Length 468;

FT	CHAIN	20	479
FT	ACT_SITE	82	62
FT	ACT_SITE	337	337
FT	DISULFID	71	387
FT	DISULFID	128	472
FT	DISULFID	216	441
FT	DISULFID	225	298
FT	DISULFID	413	421
FT	CARBOHYD	191	191
FT	CARBOHYD	315	315
FT	CARBOHYD	458	458
SQ	SEQUENCE	479 AA:	52678 MW; 4F8E0F3778CC3B08 CRC64;

Query Match
Best Local Similarity 14.0%; Score 346.5; DB 1; Length 479;
Matches 115; Conservative 54; Mismatches 160; Indels 91; Gaps

QY	49	GQYSPPFFSLADESAISDPVPGCRVTFOVLRSRGARYPTSSKSKY-SALIEAIQKNAT	107
DB	53	GYSERVS----GIARDPPTSCEVDQIVMKRGERYPSPSAGKDIEELAKVSYNTT	108
QY	108	AFKGKYAFLTKNYTL-----GADDLT-PP-GEQMVNSIGIKFYRYKAL--ARKIVPF	157
DB	109	EYKGDLAFLNDWTYYVPNECYNAETTSGGPYAGLLDAYNHGNDYKARYGHLMNGETWTFP	168
QY	158	VPRASGDRVIASAERKTGFEGPOSADLPANPHQASPVINIIPEGAGYNNTLDHGLCTA	217
DB	169	F-SSGYGVITAKRFGEF-----FGNYSTNAALNI	201
QY	218	FEESLG-----DDVEANFTAVFAPTR--ARLEAHLPGVNLTDEDVVNLMD	262
DB	202	SESEVMGADSLTPCTCDTDNDQTTCDNLTQLPKFKVAARLNSQNGMLTASDVYNLMV	261
QY	263	MCPDTVFARTSDATQLSPPCDLPTHDEWIQDYLOSLGRYKYGYGAGNPPLGPAQGVGfyne	322
DB	262	MASELNAR-----PFSNWINAFTODEWSFGVEDLVNYCYCAGPGDKNKAAVGAYANA	316
QY	323	LIALRTSHSPQDHTSTNHTLDSNPATFPFLNATLYADFSDHNTWVSIFPALGLY--NGTKP	380
DB	317	SLTLNQGPKE-----AGSLFFNFADHTNITPILAAGLVLPINEDLP	358
QY	381	LSTTSVESIETGYSASWTVPPFARAYVEMMOEA----EKEPIRLRVLVNDVRVPLHG	436
DB	359	L-----DRVAFGNPSYIGNIVPMGGHLTIERLSQATALSDSGTVRLVLEAVLPFND	413

RESULT 12
PPA5_KLUJA
ID PPA5_KLUJA STANDARD; PRT: 469 AA.

AC	P5229;
DC	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
GN	PHOS.
GC	Kluyveromyces lactis (Yeast).
OC	Eukaryota; Fungi; Ascomycota;
OX	Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
NCBI_Taxid=28985;	[1]
SEQUENCE FROM N.A.	
RC STRAIN=CBS 2359/152;	
RA Ferminan E.;	
RJ Thesis (1995);	
CC -! CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN	
CC -! SUBCELLULAR LOCATION: SECRETED.	
CC -! INDUCTION: RERESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).	
CC -! PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.	
CC -! SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collabora

Db 82 DAVGTANFQORLLGVSVPIDYVSGNPLSFVPTWTFVIEAANADALSSGRVFLDMGRQ 141
 QY 143 FYRRYKALARKIVPVRASGSDRVITASAEKEIEGQSAKLADPGANPHOASPVINVIIE 202
 Db 142 FYERHFLFNASTYNTYTAQOORVVSALWYGYGMF-----GSDVFNFTNYILVSENA 194
 QY 203 GAGYNTLDHGLCTAFESSELGDDVEANFTAVFAPPIRARLEAHLPGVNLDEDVNLMD 262
 Db 195 TAGSNLSYNACASDADDTTAPALEARNVYMPPIRQRLNPNFYSVNLNDDILLNLYG 254
 QY 263 MCPEDTAVRTSDATQSPFCDLFTHDWQIDYQSLQSGKYGYGAGNPLGPAQGVGVNE 322
 Db 255 ICSFEIALQ-----DYSEFCKLFNSVDLFNFEYEGDLSFSYGMGNSVKWGSIFGAYANS 309
 QY 323 LIARLTHSPVQDHTSTNHTLDSNPATEPLNATLYADFSHDNTVYSIFFALGLYNGTRPLS 382
 Db 310 LANSL--RSVENNTQ-----QVFFATHDANIIPVETALGFDDTNPEN 351
 QY 383 TTSVESIEETDGYAASVTPFAARAYVEMMOCEAEKEPLRVLVNDRVPLHGCG 437
 Db 352 PLPTSQVYVHSHMKASEFVFPAGNLITELFCEDSKY-YVRHLVNEEVFLPSDCG 405

RESULT 14

PPA2_SCHPO STANDARD; PRT; 463 AA.
 AC Q01682;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE THIAMINE-REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO4 OR SPEC428.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
 RX MEDLINE=91064763; PubMed=2249257;
 RA Yang J., Schweingruber M.E.;
 RT "The structural gene coding for thiamin-repressible acid phosphatase
 in Schizosaccharomyces pombe";
 RL Curr. Genet. 18:269-272(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- INDUCTION: RERESSED BY THIAMINE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X56939; CAA040258.1; --
 CC EMBL; AL034382; CAA22728.1; --
 CC PIR; S14119; S14119;
 CC InterPro; IPR000560;
 CC Pfam; PF00328; acid.phosphat.1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Cell wall; Signal.

FT SIGNAL 1 18
 FT CHAIN 19 463
 FT ACT_SITE 69 69
 FT ACT_SITE 340 340
 FT CARBOHYD 98 98
 FT CARBOHYD 104 104
 FT CARBOHYD 186 186
 FT CARBOHYD 221 221
 FT CARBOHYD 251 251
 FT CARBOHYD 328 328
 FT CARBOHYD 433 433
 FT CARBOHYD 439 439
 FT CARBOHYD 458 458
 SQ SEQUENCE 463 AA; 52118 MW; F48EAF8B8B6234A CRC64;

Query Match 12.28; Score 301; DB 1; Length 463;
 Best Local Similarity 25.48; Pred. No. 8.1e-17;

Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

QY 51 YSPFFSLADESAISPDVPKGRVTFVQVLSRHGARYPTSS-----KSKKYSALIE----- 100
 Db 43 HEYF-----NGPTSPESCAIKQVHLLQHGSRNFTGDDTATDVSSAQIDIFQNKLL 97
 QY 101 --AQKNATAFKGYAFLKTYNYTL---GADLTTPFGQQMVNSGIRYRYKALARKIV 155
 Db 98 NGSIPVNFSPENPLVYKHWTPVKAENADQLSSSGRIELFDLGRQVFRYYELFDTDV 157
 QY 156 PFVRASGSDRVITASAEKIEGFSQAKLADPGANPHOASPVINVIIE--GAGYNTLDHG 213
 Db 158 YDINTAAQERVVDSAEWFSYGM-----FGDDMNQNTF-----IVLPEDDSAGANSLAMY 208
 QY 214 LCTAFEESELGDD-VEANETA---VFAPPIRARLEAHL-PCVNLTDSDVNLMDMCPEDT 268
 Db 209 SCPVYEDNNIDENITEAHTSWRNVLKPIANRLNKYFDSGYNLTVSDVRSLYICVYEI 268
 QY 269 VARTSDATQSPFCDLFTHDWQIDYQSLQSGKYGYGAGNPLGPAQGVGVNELIARLT 328
 Db 269 ALRDN-----SDFCSLFTPSEFLNFYDSLDYAYWGGPASEWASTLGGAYVNNLANNL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATEPLNATLYADFSHDNTVYSIFFALGLYNGTKPLSTTSVES 388
 Db 323 -----RKGVNNSDRK-----VFLAFTHDSQIIPVEAALGFDPDITPHEPLPTDK 367
 QY 389 IEETDGYAASVTPFAARAYVEMMOCEAEKEPLRVLVNDRVPLHGCGVDKLG 442
 Db 368 NIFTYSLKTSFVPFAGNLITELFLC-SDNKYYVRLVNLVQVPLTDCGVPSPG 420

RESULT 15

PPAX_CAEEL STANDARD; PRT; 755 AA.
 AC Q09549; Q17843;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE ACID PHOSPHATASE P26C11.1 (EC 3.1.3.2).
 GN P26C11.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Matthews P., Lloyd C.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Sat Oct 27 15:25:41 2001

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: Z47072; CAA87370.1; -
DR EMBL: Z54342; CAA87370.1; JOINED.
DR EMBL: Z54342; CAA91156.1; -
DR EMBL: Z47072; CAA91156.1; JOINED.
DR WormPep; F26C11.1; CE05732.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein; Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
FT SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.4%; Score 134; DB 1; Length 755;
Best Local Similarity 21.9%; Pred. No. 0.0049;
Matches 91; Conservative 56; Mismatches 123; Indels 146; Gaps 23;
QY 57 LADESAISPDVPKGRVTEVQVLSRHGAR-----YPTSSSKSKYSALIEATQKNATAFK 110
Db 332 LAIEKSFMEV-DNMKLEFVQIWRHGDRSALEGLFPISEK----- 371
QY 111 GKYAFLEKTYNYLGA---DDLTPFGQOMVNSGIKFYRY-----KALARKIVPEV 158
Db 372 -----NWTFGGGGLGELTPMGSEMNNLGTIFRRRYVEDQOFLSHRYAAKEI--YI 420
QY 159 RASGSDRVIASAEKFTGFSQAKLADPGANPHQASPVINVIIEGAGYNNITLD----- 211
Db 421 RSTNLNLTIIISAMSLYG-----MFPFGA-----WNI---QGVDPNDVDWQOGFTF 464
QY 212 -----HGL-----CTAFEE-----SELGDDVEANFTAVFAPPPIRAREAHLPGVN 251
Db 465 IPHVVDGIDQCAVACLCNCRFRFOELQEKWAEI-DEYKNATVAMIA--LNRVAAF---YN 518
QY 252 LTDEDVNLMDMCFDFRVARTSDA--TQLSPFCD-LFTHDEWIOYDYLQSLGKYG---- 304
Db 519 VTQDP-----EKFNRYTDANKCQRNWFNDTWYQQLPWYNEDLYNEAQTYAPEKR 568
QY 305 ---YGAGNP-----LGPAGGVGFVNELIARLTHSPVQDHTSTNHTLD--SNPA 347
Db 569 FTEGFGNPKPSIVDGDIPQEVSTLQGGPLLNEIFERGKIRCVADAENCSIDYIKPL 628
QY 348 TFPLNATLYADFSDHNTWVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPF 403
Db 629 KF-----YAYSSHDLQVYALLVTLGI-----TDVKTVDGWPDTSSSLTIEY 670

Search completed: October 26, 2001, 15:18:02
Job time: 448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: October 26, 2001, 15:18:09 ; Search time 32.57 Seconds
(without alignments)
295.232 Million cell updates/sec

Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGVPVLLSIATLFGSTSGT.....DFVEGLSFARSGGNWEECPA 467

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/6C_COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2185	88.5	441	4	US-09-121-425-1
2	2173	88.0	467	4	US-09-121-425-2
3	1866	75.5	467	1	US-07-923-724-8
4	1866	75.5	467	2	US-08-609-426A-8
5	1866	75.5	467	2	US-08-374-652C-2
6	1862	75.4	467	1	US-08-151-574-32
7	1862	75.4	467	1	US-08-146-424-20
8	1862	75.4	467	1	US-08-693-709-2
9	1862	75.4	467	2	US-08-419-448-32
10	1862	75.4	467	2	US-08-819-825-3
11	1855	75.1	467	4	US-09-163-642-3
12	1813	73.4	443	4	US-09-155-855-3
13	1808	73.2	443	4	US-09-155-855-1
14	1395	56.5	475	2	US-08-155-855-2
15	1395	56.5	475	2	US-08-155-855-2
16	846.5	34.3	443	3	US-08-993-359-30
17	846.5	33.7	453	3	US-08-993-359-22
18	831.5	33.7	439	3	US-08-993-359-24
19	831.5	33.7	439	3	US-09-221-654-2
20	831.5	33.7	439	3	US-08-989-358A-2
21	822	33.3	442	3	US-08-993-359-28
22	811	32.8	442	3	US-08-993-359-26
23	359	14.5	468	1	US-07-627-539G-2
24	354.5	14.4	446	1	US-07-627-539G-7
25	346.5	14.0	479	1	US-07-923-724-2
26	346.5	14.0	479	2	US-08-609-426A-2
27	346.5	14.0	479	2	US-08-609-426A-2

28	346.5	14.0	479	2	US-08-374-652C-4
29	185.5	7.5	92	3	US-08-993-359-32
30	179	7.2	449	3	US-08-680-506-7
31	121.5	4.9	318	3	US-08-680-506-3
32	121	4.9	113	1	US-08-241-853-8
33	121	4.9	113	1	US-08-850-917-8
34	115	4.7	24	2	US-08-374-652C-32
35	110	4.5	113	1	US-08-241-853-10
36	110	4.5	113	2	US-08-850-917-10
37	98	4.0	386	1	US-08-758-213-1
38	98	4.0	386	2	US-08-692-787-48
39	98	4.0	386	4	US-09-097-199-48
40	98	4.0	515	2	US-09-146-283-2
41	98	4.0	515	3	US-08-579-823A-2
42	98	4.0	515	4	US-09-344-195-2
43	97	3.9	35	1	US-08-151-574-2
44	97	3.9	35	2	US-08-419-448-2
45	93.5	3.8	675	4	US-09-171-878-1

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
US-09-121-425-1

Query Match	Best Local Similarity	Score	2185;	DB 4;	Length 441;
Matches	409;	Conservative	13;	Mismatches	19;
Indels	0;	Gaps	0;		
QY	27	NSHSCDVTGQYQCFPEISHLWGQYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY	86		
Db	1	NSHSCDVTGQYQCFPEISHLWGQYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY	60		
QY	87	PTSKSKKYSALIEAIOKNAFAFGYAFKTYNYTLGADDLTPFGEQOMVNSGIFKFR	146		
Db	61	PTSKSKKYSALIEAIOKNAFAFGYAFKTYNYTLGADDLTPFGEQOMVNSGIFKFR	120		
QY	147	YKALARKIVPFVPSASSDRVIAAEKFIQFQSAKLADPGANPHOASPVINVIIPEGAGY	206		
Db	121	YKALARKIVPFVPSASSDRVIAAEKFIQFQSAKLADPGANPHOASPVINVIIPEGAGY	180		
QY	207	NNTLDHGLCTAFESSELGDDVEANFTAFAPIRARLEAHLPGVNLTDDEVVNLMDMCP	266		
Db	181	NNTLDHGLCTAFESSELGDDVEANFTAFAPIRARLEAHLPGVNLTDDEVVNLMDMCP	240		
QY	267	DTVARTSDATQLSPDCDLFTHDEWTQYDYLQSLGKYYGYGAGNPLGPAQGVGFNELLAR	326		
Db	241	ETVARTSDATQLSPDCDLFTHDEWTQYDYLQSLGKYYGYGAGNPLGPAQGVGFNELLAR	300		
QY	327	LTHSPVDHTNTNTHLDSNPATFPLNATLYADFSHDNTMVSIFPFGALGYNKTKPLSTTSV	386		

Sat Oct 27 15:25:39 2001

Db 301 LTRSPVQDHTSTNHTLDSNPATFFPLNLTLYADFSHDSNMISIFFALGLYNGTAPLSTTSV 360
QY 387 ESIEETDGYAASWTVPFAARAYVEMMOCEAEKPLRVLVNDRVPLHGGVKGKGRKR 446
Db 361 ESIEETDGYASWTVPFCARAYVEMMOCEAEKPLRVLVNDRVPLHGGCAVDKLGCKR 420
QY 447 DDFVEGLSFARSGNWECEFA 467
Db 421 DDFVEGLSFARSGNWECEFA 441
RESULT 2
US-09-121-425-2
; Sequence 2, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2
Query Match 88.0%; Score 2173; DB 4; Length 467;
Best Local Similarity 85.2%; Pred. No. 7.3e-221; Indels 40; Gaps 2;
Matches 415; Conservative 13; Mismatches 19;
QY 1 MGVEVLLSTATLFGSTGALTALPGRNHSCDVTGGYQCFPEISHLMGQYSPFFSLADE 60
Db 1 MGVEVLLSTATLFGSTGALTALPGRNHSCDVTGGYQCFPEISHLMGQYSPFFSLEDE 60
QY 61 SAISPDVPKGCRTVVOVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPKGCRTVVOVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
QY 121 YTLGADDLTPGEOMVNSGKIFVRRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
Db 101 YTLGADDLTPGEOMVNSGKIFVRRYKALARKIVPFIRASGSDRVIAAEKFIQFQSA 160
QY 181 KLADPGANPHQASPVIN-----VITPEGAGYNNTLDHGLCTAFEE 220
Db 161 KLADPGSOPHQASPVVIDLIEAIQKNATAFKGYAFLKVIIPESGYNNTLDHGLCTAFED 220
QY 221 SELGDDVEANFTAFAPPIRLEAHLPGVNLTDDEDVNLMDMCPPTVARTSDATQLSP 280
Db 221 SELGDDVEANFTAFAPPIRLEADLPGLVTLTDEDVNLMDMCPPTVARTSDATLSP 280
QY 281 FCDLFTHDWIOYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLTHSPVDHTSTNH 340
Db 281 FCALFTHDWIRYDYLQSLGKYGYGAGNPLGPAQGVGFANELLARLTHSPVDHTSTNH 340
QY 341 TLDNSNPATPLNATLYADFSHDSNMISIFFALGLYNGTAPLSTTSVSEIETDGYASWT 400
Db 341 TLDNSNPATPLNATLYADFSHDSNMISIFFALGLYNGTAPLSTTSVSEIETDGYASWT 400
QY 401 VPFAARAYVEMMOCEAEKPLRVLVNDRVPLHGGVKGKGRKDDFVEGLSFARSGG 460
Db 401 VPFCARAYVEMMOCEAEKPLRVLVNDRVPLHGGCAVDKLGCKRDDFVEGLSFARSGG 460
QY 461 NWECEFA 467
Db 461 NWECEFA 467

Db 461 NWECEFA 467
RESULT 3
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambossek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8
Query Match 75.5%; Score 1866; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 1.9e-188; Indels 0; Gaps 0;
Matches 348; Conservative
QY 1 MGVEVLLSTATLFGSTGALTALPGRNHSCDVTGGYQCFPEISHLMGQYSPFFSLADE 60
Db 1 MGVEVLLPLYLAGVTSLGAVPASRNOSTCDTVDQGYQCFSETHLWGOYAPFFSLANE 60
QY 61 SAISPDVPKGCRTVVOVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPKGCRTVVOVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120


```

QY 121 YTLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKTFIEGFQSA 180
Db 121 YSLGADLLTPFGEQELVNSGKIFRYVESLTRNLIPIRSGSSRVIASGEKTFIEGFQST 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVISEASSNNTLDPCGTCTVEFSELDADVEANFTATFAPSIR 240
QY 241 ARLEAHLPGVNLTDDEVDVNLDMCPDFTVARTSDATQSPFCDLTFHDEWIDYDLQSLG 300
Db 241 ORLENDLSGVTLDTEVTYLMDCSFDITSTVDTKLSPFCDLTFHDEWIDYDLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGPTQGVGNELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HNTMVSIFFALGLXNGTKPLSTTSVESTTEEDGYAAASWTVPFAARAYVEMMOCEAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTVENITQDGFSSAWTVPFASRLYVEMMOCEAKEP 420
QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467
Db 421 LVRVLNDRVPLHGGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

```

RESULT 4

```

US-08-609-426A-8
; Sequence 8, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.

```

```

; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-426A-8

```

```

Query Match 75.5%; Score 1866; DB 2: Length 467;
Best Local Similarity 74.5%; Pred. No. 1.9e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
QY 1 MGVEVLLSTATLFGSTGTFALGPRGNSHSDTVGGYQCFPEISHLWGQYSPFFSLADE 60
Db 1 MGYSAVLLPYLLAGVTSGLAVPASRNQSTCDTVDOGYQCFSETSHLWGQYAPFFSLANE 60
QY 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSKKYSALIEATOKNATAFKGYAFKLTYN 120
Db 61 SAISPDVPGCRVTFQAQLSRHGARYPTESKGYKYSALIEEQNNVTFDGYAFKLTYN 120
QY 121 YTLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKTFIEGFQSA 180
Db 121 YSLGADLLTPFGEQELVNSGKIFRYVESLTRNLIPIRSGSSRVIASGEKTFIEGFQST 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVISEASSNNTLDPCGTCTVEFSELDADVEANFTATFAPSIR 240
QY 241 ARLEAHLPGVNLTDDEVDVNLDMCPDFTVARTSDATQSPFCDLTFHDEWIDYDLQSLG 300
Db 241 ORLENDLSGVTLDTEVTYLMDCSFDITSTVDTKLSPFCDLTFHDEWIDYDLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGPTQGVGNELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HNTMVSIFFALGLYNGTKPLSTTSVESTTEEDGYAAASWTVPFAARAYVEMMOCEAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTVENITQDGFSSAWTVPFASRLYVEMMOCEAKEP 420
QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467
Db 421 LVRVLNDRVPLHGGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

```

RESULT 5

```

US-08-374-652C-2
; Sequence 2, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; NUMBER OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON

```

us-09-488-265-26.ra1

Sat Oct 27 15:25:39 2001

```

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: PCT/US93/07058
APPLICATION NUMBER: 27-JUL-1993
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-374-652C-2

```

```

Query Match 75.4%; Score 1866; DB 2; Length 467;
Best Local Similarity 74.5%; Pred. No. 1.9e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
QY 1 MGVFVLLSIATLFGSTGATLGRGNHSCDVTVDGYQCFPEISHLWQYSPFFSLADE 60
DB 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSCDVTVDGYQCFSETSHLWQYAPFFSLANE 60
QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFKLTYN 120
DB 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFKLTYN 120
QY 121 YTLGADDLTPFGQOMVNSGKIFKRYRKYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFKRYRKYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
QY 181 KLADPGANPHQASPVINIIPEGAGYNNLTDLHGCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGANPHQASPVINIIPEGAGYNNLTDLHGCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWTOYDYLQSLG 300
DB 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWTOYDYLQSLG 300
QY 301 KYGCGAGNPLGPGQGVGNELIARLTHSPVQDHTSNHLDNSNPATFPLNATLYADF 360
DB 301 KYGCGAGNPLGPGQGVGNELIARLTHSPVQDHTSNHLDNSNPATFPLNATLYADF 360
QY 361 HDNMTSIFPALGLYNGTKPLSTTSVESIEETDGYAASWTVPFARAYVENMQCEAEK 420
DB 361 HDNMTSIFPALGLYNGTKPLSTTSVESIEETDGYAASWTVPFARAYVENMQCEAEK 420
QY 421 LVRVLNDRVVPVHGGGVDRKLGCRKRDDEVEGLSFARSGNNEECPA 467
DB 421 LVRVLNDRVVPVHGGGVDRKLGCRKRDDEVEGLSFARSGNNEECPA 467

```

```

RESULT 6
US-08-151-574-32
Sequence 32, Application US/08151574
Patent No. 5436156
GENERAL INFORMATION:
APPLICANT: Robert F.M. Van Gorcom
APPLICANT: Willem Van Hartingsveldt
APPLICANT: Petrus A. Van Paridon
APPLICANT: Annemarie E. Veenstra
APPLICANT: Rudolf G.M. Luttin
APPLICANT: Gerardus Selten
TITLE OF INVENTION: Cloning and Expression of Microbial
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025-3471
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/151,574
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/688,578
FILING DATE: 24-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20026.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-151-574-32

```

```

Query Match 75.4%; Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 5.2e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
QY 1 MGVFVLLSIATLFGSTGATLGRGNHSCDVTVDGYQCFPEISHLWQYSPFFSLADE 60
DB 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSCDVTVDGYQCFSETSHLWQYAPFFSLANE 60
QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFKLTYN 120
DB 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFKLTYN 120
QY 121 YTLGADDLTPFGQOMVNSGKIFKRYRKYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFKRYRKYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
QY 181 KLADPGANPHQASPVINIIPEGAGYNNLTDLHGCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGANPHQASPVINIIPEGAGYNNLTDLHGCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWTOYDYLQSLG 300
DB 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWTOYDYLQSLG 300

```

QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 DB 301 KYGYGAGNPLGPTGGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HNTWVSIFFAFGALYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGALYNGTKPLSTTTVENITQDGFSSAWTPFASRLYVEMMOCEAEQEP 420
 QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 DB 421 LVRVLVNDRVVPLHGCPVDALGRCTRDSFVRGLSFARSGDWAECFA 467
 RESULT 7
 US-08-146-424-20
 ; Sequence 20, Application US/08146424
 ; Patent No. 5593963
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/146,424
 ; FILING DATE: 02-NOV-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KENNEDY, BILL
 ; REGISTRATION NUMBER: 33,407
 ; REFERENCE/DOCKET NUMBER: 44615-20011.24
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-146-424-20
 Query Match 75.4%; Score 1862; DB 1; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
 QY 1 MGTVFVLLSINATLFGSTSGTALGRGNHSCDVTVDGQYQCFPELISHWQYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLVAPASRNQSCDITVDQYQCFSETSHLWQYAPFFSLANE 60
 QY 61 SATSPDVPKCRVTFVQVLSHRGARYPTSSKKYSALIEAIQKNATAFKGYAEFLKTN 120
 DB 61 SVTSPVPACRVTFACVLSHRGARYPTSDKGYKYSALIEEQQNATTFDGYAEFLKTN 120
 QY 121 YTLGADDLTPFGEGQMVNSGINKFYRRYKALARKIVFVRASGSDRVIASAEKFIQFQSA 180

DB 121 YSLGADDLTPFGCEQLVNSGINKFYRYESLRTNIVFIRSSGSSRVIASGKKFIEGFQST 180
 QY 181 KLADPCANPHQASPVNVIIPGAGYNNNTLDHGLCTAFEESELGDDVEANFTAVFAPDIR 240
 DB 181 KLDPRAPQGGSSPKIDVVISSESSNNTLDPGTCTVFEDSELADTVEANFTAVFVPSIR 240
 QY 241 ARLEAHLFGVNLTDREVVNLMDMCPFDITVARTSDATQLSPFCDLTFHDEWIOYDYLOSGL 300
 DB 241 QRLNDLSGVTLTDTDEVYLMDCSFDITSTVDTKLSPFCDLTFHDEWINYDYLOSGL 300
 QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 DB 301 KYGYGAGNPLGPTGGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HNTWVSIFFAFGALYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGALYNGTKPLSTTTVENITQDGFSSAWTPFASRLYVEMMOCEAEQEP 420
 QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 DB 421 LVRVLVNDRVVPLHGCPVDALGRCTRDSFVRGLSFARSGDWAECFA 467

RESULT 8
 US-08-693-709-2
 ; Sequence 2, Application US/08693709
 ; Patent No. 5770413
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/693,709
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146,424
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20011.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-813-5600
 ; TELEFAX: 415-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal

us-09-488-265-26.ra1

Sat Oct 27 15:25:39 2001

```

FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1...23
; OTHER INFORMATION:
US-08-693-709-2

Query Match      75.4%  Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 5.2e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVSAYLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCFPEITSHLWGOYSPFFSLADE 60
Db 1 MGVSAYLLPLLYLLSGVTSLGAVPASRNQSSCDTVDQGYQCFSEITSHLWGOYAPFFSLANE 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGQFSA 180
Db 121 YSLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGQFSA 180

QY 181 KLADPGANPHQASPVNVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240
Db 181 KLADPGANPHQASPVNVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240

QY 181 KLKDPRAQPGOSSPKIDVISEASSNNITLDPGCTCTVFEDSELADTVEANFTAVFAPP 240
Db 181 KLKDPRAQPGOSSPKIDVISEASSNNITLDPGCTCTVFEDSELADTVEANFTAVFAPP 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDFLTHDWDIYDLSLG 300
Db 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDFLTHDWDIYDLSLG 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHSTNHTLDSNPATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHSTNHTLDSNPATPLNATLYADES 360

QY 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAKEP 420
Db 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAKEP 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVGLSFARSGGNWECFA 467
Db 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVGLSFARSGGNWECFA 467

RESULT 9
US-08-419-448-32
; Sequence 32, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorkom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Parijs
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448

```

```

FILING DATE: 10-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-419-448-32

Query Match      75.4%  Score 1862; DB 2; Length 467;
Best Local Similarity 74.5%; Pred. No. 5.2e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVSAYLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCFPEITSHLWGOYSPFFSLADE 60
Db 1 MGVSAYLLPLLYLLSGVTSLGAVPASRNQSSCDTVDQGYQCFSEITSHLWGOYAPFFSLANE 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGQFSA 180
Db 121 YSLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTIEGQFSA 180

QY 181 KLADPGANPHQASPVNVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240
Db 181 KLADPGANPHQASPVNVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240

QY 181 KLKDPRAQPGOSSPKIDVISEASSNNITLDPGCTCTVFEDSELADTVEANFTAVFAPP 240
Db 181 KLKDPRAQPGOSSPKIDVISEASSNNITLDPGCTCTVFEDSELADTVEANFTAVFAPP 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDFLTHDWDIYDLSLG 300
Db 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDFLTHDWDIYDLSLG 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHSTNHTLDSNPATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHSTNHTLDSNPATPLNATLYADES 360

QY 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAKEP 420
Db 361 HDNTMVSIFFAFGLYNGTKPLSTTSVESIETDGYAASWTPFAARAYVEMMOCEAKEP 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVGLSFARSGGNWECFA 467
Db 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVGLSFARSGGNWECFA 467

RESULT 10
US-08-819-825-3
; Sequence 3, Application US/08819825
; Patent No. 5866118
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,825
 FILING DATE: 18-MAR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4758.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-819-825-3

Query Match 75.4%; Score 1862; DB 2; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGCVFVLLSIATLFGSTGALGPRGNHSCDVTVDGQYQCFPEISHLWGOYSPFFSLADE 60
 Db 1 MGCVFVLLSIATLFGSTGALGPRGNHSCDVTVDGQYQCFPEISHLWGOYSPFFSLADE 60
 QY 61 SAISPDPKGRVTFVQVLSRHGARYPTSSKSKYKYSALIEAIQKNATFAFKYAFKLTYN 120
 Db 61 SVISPEVPAGCRVTFQVLSRHGARYPTSSKSKYKYSALIEAIQKNATFAFKYAFKLTYN 120
 QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 Db 121 YSLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 QY 181 KLADPGANPHQASPVINVIPEAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPGANPHQASPVINVIPEAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIQDYLSLG 300
 Db 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIQDYLSLG 300
 QY 301 KYGYGAGNPLGPTQGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNLTLYADFS 360
 Db 301 KYGYGAGNPLGPTQGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNLTLYADFS 360
 QY 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCAEKEP 420
 Db 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCAEKEP 420
 QY 421 LVRVLNDRVPLHCGCPVDALGRCTRODSFVRGLSFARSGGDWAECEFA 467
 Db 421 LVRVLNDRVPLHCGCPVDALGRCTRODSFVRGLSFARSGGDWAECEFA 467

RESULT 11
 US-09-163-642-3
 ; Sequence 3, Application US/09163642
 ; Patent No. 6221644
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Ray, Michael W.
 ; APPLICANT: Klotz, Alan V.
 ; TITLE OF INVENTION: Polypeptides Having Phytase Activity
 ; TITLE OF INVENTION: And Nucleic Acids Encoding Same
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 62216440 No. 6221644disk of No. 6221644th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/163,642
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,825
 FILING DATE: 18-MAR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4758.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-163-642-3

Query Match 75.4%; Score 1862; DB 4; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGCVFVLLSIATLFGSTGALGPRGNHSCDVTVDGQYQCFPEISHLWGOYSPFFSLADE 60
 Db 1 MGCVFVLLSIATLFGSTGALGPRGNHSCDVTVDGQYQCFPEISHLWGOYSPFFSLADE 60
 QY 61 SAISPDPKGRVTFVQVLSRHGARYPTSSKSKYKYSALIEAIQKNATFAFKYAFKLTYN 120
 Db 61 SVISPEVPAGCRVTFQVLSRHGARYPTSSKSKYKYSALIEAIQKNATFAFKYAFKLTYN 120
 QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 Db 121 YSLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 QY 181 KLADPGANPHQASPVINVIPEAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPGANPHQASPVINVIPEAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIQDYLSLG 300
 Db 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIQDYLSLG 300
 QY 301 KYGYGAGNPLGPTQGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNLTLYADFS 360
 Db 301 KYGYGAGNPLGPTQGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNLTLYADFS 360
 QY 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCAEKEP 420
 Db 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCAEKEP 420
 QY 421 LVRVLNDRVPLHCGCPVDALGRCTRODSFVRGLSFARSGGDWAECEFA 467
 Db 421 LVRVLNDRVPLHCGCPVDALGRCTRODSFVRGLSFARSGGDWAECEFA 467

RESULT 12
 US-09-155-855-3

us-09-488-265-26.ra

Sat Oct 27 15:25:39 2001

FILE REFERENCE: 81356/124
 CURRENT APPLICATION NUMBER: US/09/155,855
 CURRENT FILING DATE: 1998-10-05
 EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
 EARLIER FILING DATE: 1997-04-04
 EARLIER APPLICATION NUMBER: JP 084314
 EARLIER FILING DATE: 1996-04-05
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 443
 TYPE: PRT
 ORGANISM: Aspergillus niger
 US-09-155-855-1

Query Match 73.4%; Score 1813; DB 4; Length 443;
 Best Local Similarity 75.1%; Pred. No. 7.1e-183; Indels 0; Gaps 0;
 Matches 331; Conservative 49; Mismatches 61;
 QY 27 NNSHCDTVDGGYQCFPEJSHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 86
 DB 3 NQSTCDTVDGGYQCFSETSHLWGOYAPFFSLANKSAISPDVPGCHVTFAQVLSRHGARY 62
 QY 87 PTSSKSKYSALIEAIOKNATAFKGYAFLKTYNTLGCADDLTPEGEOMVNSGKYYRR 146
 DB 63 PTDSKSKYSALIEEIQONATTPEGYAFLKTYNTSLGADDLTPEGEOLVNSGKYYR 122
 QY 147 YKALARKIYVPRASGSDRVASAEKFIQFQSAKLADPGANPHQASPVINIIPEGAGY 206
 DB 123 YESLTRNIVPFIKSSSSRSVIASGNKFIQFQSTKLKDPRAQGGSPKIDVISEASTS 182
 QY 207 NNTLDHGLCTAFSESELGDDVEANFTAVFAPIRAPLEAHLPGVNLTDDEVVNLMDMCP 266
 DB 183 NNTLDHGLCTAFSESELGDDVEANFTAVFAPIRAPLEAHLPGVNLTDDEVVNLMDMCP 242
 QY 267 DTVARTSDATQSPFCDLTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGVNELIAR 326
 DB 243 DTISTVTDTKLSPFCDLTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGVNELIAR 302
 QY 327 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFPAGLYNGTKPLSTTSV 386
 DB 303 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFPAGLYNGTKPLSTTSV 362
 QY 387 ESIEETDCYAAASWVPPFAARAYVEMMOCEAKEPLVRVLVNDVRVPLHGGGVDKLGRCKR 446
 DB 363 ENITQDGFSSAWVPPFAARAYVEMMOCEAKEPLVRVLVNDVRVPLHGGGVDKLGRCKR 422
 QY 447 DDFVEGLSFARSGGNWEECPA 467
 DB 423 DDFVEGLSFARSGGNWEECPA 443

RESULT 14
 US-09-155-855-2
 Sequence 2, Application US/09155855
 Patent No. 6139902
 GENERAL INFORMATION:
 APPLICANT: KONDO, Hidemasa
 APPLICANT: ANAZAWA, Hideharu
 APPLICANT: KANEKO, Syunichi
 APPLICANT: NAGASHIMA, Tadashi
 APPLICANT: TANGE, Tatsuya
 TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
 FILE REFERENCE: 81356/124
 CURRENT APPLICATION NUMBER: US/09/155,855
 CURRENT FILING DATE: 1998-10-05
 EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
 EARLIER FILING DATE: 1997-04-04
 EARLIER APPLICATION NUMBER: JP 084314
 EARLIER FILING DATE: 1996-04-05
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.0

Sequence 3, Application US/09155855
 Patent No. 6139902
 GENERAL INFORMATION:
 APPLICANT: KONDO, Hidemasa
 APPLICANT: ANAZAWA, Hideharu
 APPLICANT: KANEKO, Syunichi
 APPLICANT: NAGASHIMA, Tadashi
 APPLICANT: TANGE, Tatsuya
 TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
 FILE REFERENCE: 81356/124
 CURRENT APPLICATION NUMBER: US/09/155,855
 CURRENT FILING DATE: 1998-10-05
 EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
 EARLIER FILING DATE: 1997-04-04
 EARLIER APPLICATION NUMBER: JP 084314
 EARLIER FILING DATE: 1996-04-05
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 467
 TYPE: PRT
 ORGANISM: Aspergillus niger
 US-09-155-855-3

Query Match 75.1%; Score 1855; DB 4; Length 467;
 Best Local Similarity 73.4%; Pred. No. 2.8e-187; Indels 0; Gaps 0;
 Matches 343; Conservative 51; Mismatches 73;
 QY 1 MGVFVLLSIATLFGSTGALPRGNHSCDVTVDGGYQCFPEIHLWGOYSPFFSLADE 60
 DB 1 MGVSALLPLVLLSGVTSGLAVPASRQSTCDTVDGGYQCFSETSHLWGOYAPFFSLANK 60
 QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFLKTYN 120
 DB 61 SAISPDVPGCHVTFAQVLSRHGARYPTDSKSKYSALIEEIQONATTPEGYAFLKTYN 120
 QY 121 YTLGADDLTPFEQOMVNSGKYYRRYKALARKIYVPRASGSDRVASAEKFIQFQSA 180
 DB 121 YSLGADDLTPFEQOMVNSGKYYRRYKALARKIYVPRASGSDRVASAEKFIQFQSA 180
 QY 181 KLADPGANPHQASPVINIIPEGAGYNTLDPGLCTAFSESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQGGSPKIDVISEASTSNLDPGCTVFEDELADDEANFTAVFAPPIR 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSPFCDLTHDEWIQYDYLQSLG 300
 DB 241 QRLNGLSGVSLTDETVYLMDCSFDTISTVTDTKLSPFCDLTHDEWIQYDYLQSLN 300
 QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPTQGVGVANIELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 QY 361 HDNTWYSIFPAGLYNGTKPLSTTSVESIEETDCYAAASWVPPFAARAYVEMMOCEAKEP 420
 DB 361 HDNGIILFALGLYNGTKPLSTTSVAENITQDGFSSAWVPPFAARAYVEMMOCEAKEP 420
 QY 421 LVRVLVNDVRVPLHGGGVDKLGRCKRDDFVEGLSFARSGGNWEECPA 467
 DB 421 LVRVLVNDVRVPLHGGGVDKLGRCKRDDFVEGLSFARSGGNWEECPA 467

RESULT 13
 US-09-155-855-1
 Sequence 1, Application US/09155855
 Patent No. 6139902
 GENERAL INFORMATION:
 APPLICANT: KONDO, Hidemasa
 APPLICANT: ANAZAWA, Hideharu
 APPLICANT: KANEKO, Syunichi
 APPLICANT: NAGASHIMA, Tadashi
 APPLICANT: TANGE, Tatsuya
 TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

```

; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown
US-09-155-855-2

```

```

Query Match
Best Local Similarity 73.2%; Score 1808; DB 4; Length 443;
Matches 330; Conservative 49; Mismatches 59; Indels 0; Gaps 0;
QY 30 SCDTVGGYOCFPEISHLWGQYSPFFSLADESAISPDPKGRVTFVQVLSRHGARYPTS 89
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 TCDTVGGYOCFSETSHLWGQYAPFFSLANKSAISPDPAGCHVTFPAQVLSRHGARYPD 65
QY 90 SKSKKYSALTEAOKNATKPKYAEKLYNTYTLGADDLTPFGEQOMVNSGKFKYRKYA 149
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 SKGKYSALIEEQNATTFEGKYAEKLYNTYSLGADDLTPFGEQELVNSGVKFYQRYES 125
QY 150 LAKKIVFVRASGSDRVIAAEKFIKGFOSAKIADPGANPHOASPVINVIPEGAGYNT 209
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 LTRNIIVFTRSSGSSRVIASGNKFIKGFQSTKLKDPRAQPGSSPKRIDVISEASTSNT 185
QY 210 LDHGLCTAFESEELGDDVFNATVAPPFIRARLEAHLPGVNLTDDEVVNLMDMCPFDTV 269
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 LDPGCTVFESELDADDEANFATFVPSIRORLENDLSGVSLTDEVTYLMDCSFDTI 245
QY 270 ARTSDATQSPCEDLFTHDEWIOYDLOSLGKYGYGAGNPLGPAQGVGFVNELIARLTH 329
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 STSTVDTKLSPCEDLFTHEWINDYDLOSLKYYGHGAGNPLGTPQGVGYANELIARLTH 305
QY 330 SPVQDHTSNHTLDSNPATPLNATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSVESI 389
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 SPVHDTSSNHTLDSNPATPLNSTLYADFSDHNGIISILFALGLYNGTKPLSTTAEI 365
QY 390 EETDGYAASWTVPFAARAYVEMMOCEAEKEPLRVLVNDRVVPVPLHGCVDKLGRCRDRDF 449
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 TQDGFSSAWTVFASRMVYEMMOCSOEPLRVLVNDRVVPVPLHGCVDKLGRCRDRDF 425
QY 450 VEGLSFARSGGNWEECEFA 467
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
426 VKGLSFARSGGDWGECEFA 443

```

```

RESULT 15
US-08-819-825-2
; Sequence 2, Application us/08819825
; Patent No. 5866118
; GENERAL INFORMATION:
; APPLICANT: Berkay, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5866118 of No. 5866118 disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-819-825-2

```

```

Query Match
Best Local Similarity 56.5%; Score 1395; DB 2; Length 475;
Matches 285; Conservative 59; Mismatches 105; Indels 38; Gaps 11;
QY 1 MGVP-VVLLSTATLFGSTSGTALGP--RGNSHSCTVDGGYQCPEISHLWGOVSPFESL 57
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 LGSELVLLLOFSALL--TASPAIPFPFWRKHPNVD-----IARHWGOYSFESL 52
QY 58 ADESAISPDPVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAOKNATKPKYAEKFI 117
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 AEVSEISPAPVPGKCRVTFVQVLSRHGARYPTAHKSEYAELOQIOTATFEFGDEAFRL 112
QY 118 TYNITGLADLTPFGEQOMVNSGKFKYRKYKALKARKIVPFVRASGSDRVIASAEKFI 177
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 DYAYHLGADNLTRFGEQOMESGRQFVHYRQAREIVFVRASGARSARVIAAEFFNRGF 172
QY 178 QSAKLADPGANPHOASPVINVIPEGAGYNTLDHGLCTAFESEELGDDVE--ANFTAVFA 236
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 QDAKDRDPSNKDOAEPIVINIISSETGSNNLTDGLTCCPAAEAP--DPTQPAEFLQVFG 230
QY 237 PPIRABLEAHLPGVNLTDDEVVNLMDMCPFDTVARTSDAT----QLSPCEDLFTHDEWIO 292
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 PRVKKITKHPGVNLTDDEVVPLEMDLCPFDIVG--SDPVLFPRLSPCHLFTADDNWA 288
QY 293 YDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLT-HSPVQDHTSNHTLDSNPATEPL 351
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 YDYTYTLDRYISHGGSAGFSGRGVGFVNELIARMTGNLPVKDHTTVNHTLDDNPETPL 348
QY 352 NATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSVE--SIBETDGYAASWTVPFAARAYV 409
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 DAVLYADFSDHNTMTGIFSAMGLYNGTKPLSTSKIQPTGAAADGYAASWTVPFAARAYV 408
QY 410 ENMOC-----EAEKEPLRVLVNDRVVPVPLHGCVDKLGRCRDRDFVEGLSFARSG 459
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
460 GNWEECF 466
QY 469 GHWDRCF 475
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: October 26, 2001, 16:40:45
Job time: 4956 sec

us-09-488-265-26.ra1

Sat Oct 27 15:25:39 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:29 ; Search time 57.24 Seconds
(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGFEVLLSIATLFGSTSGT.....DFVEGLSPARSGGNWEECA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2470	100.0	467	21	Consensus phytase
2	2470	100.0	467	21	Phytase-10, a cons
3	2448	99.1	467	20	Consensus phytase
4	2448	99.1	467	21	Consensus phytase
5	2448	99.1	467	21	Mutant phytase-10,
6	2442	98.9	467	21	Consensus phytase
7	2436	98.6	467	21	Consensus phytase
8	2353	95.3	467	21	Consensus phytase
9	2347	95.0	467	21	Consensus phytase
10	2344	94.9	441	21	Consensus phytase
11	2342	94.8	467	20	Consensus phytase

12	2342	94.8	467	21	Consensus phytase
13	2342	94.8	467	21	Mutant phytase-1,
14	2311	93.6	467	20	Fungal phytase pro
15	2311	93.6	467	21	Consensus phytase
16	2311	93.6	467	21	Phytase-1, a conse
17	2305	93.3	467	20	Fungal consensus p
18	2304	93.3	467	20	Fungal consensus p
19	2304	93.3	467	20	Ascomycete consens
20	2296	93.2	467	20	Fungal consensus p
21	2296	93.0	467	20	Fungal consensus p
22	2295	92.9	467	20	Consensus phytase
23	2291	92.8	467	21	Consensus phytase
24	2291	92.8	467	21	Phytase-7, a deriv
25	2204	89.2	467	20	Consensus phytase
26	2204	89.2	467	21	Consensus phytase
27	2201	89.1	467	21	Phytase-11, a cons
28	2185	88.5	441	21	Consensus phytase
29	2115	85.6	437	21	Consensus phytase
30	2115	85.6	437	21	Phytase-11, a cons
31	2066	83.6	431	21	Consensus phytase
32	2066	83.6	431	21	Initial consensus
33	2050.5	83.0	424	21	Consensus phytase
34	2050.5	83.0	424	21	Consensus phytase
35	2048	82.9	467	21	Initial consensus
36	2028	82.1	467	20	Mutant Aspergillus
37	2028	82.1	467	21	A. fumigatus phyt
38	2028	82.1	467	21	Aspergillus fumiga
39	1902	77.0	467	21	Mutant Aspergillus
40	1902	77.0	465	19	Aspergillus fumiga
41	1892.5	76.6	440	21	A. fumigatus phyt
42	1889.5	76.5	440	21	Aspergillus fumiga
43	1889.5	76.5	440	21	Aspergillus fumiga
44	1884.5	76.3	440	21	Aspergillus fumiga
45	1884.5	76.3	440	21	Aspergillus fumiga

ALIGNMENTS

RESULT 1
AAB20524
ID AAB20524 standard; Protein; 467 AA.
XX
AC AAB20524;
XX
DT 05-DEC-2000 (first entry)
XX
DE Consensus phytase 10 SEQ ID NO:26.
XX
KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure.
XX
OS Synthetic.
XX
PN WO200043503-A1.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-DK00025.
XX
PR 22-JAN-1999; 99DK-0000092.
XX
PR 21-SEP-1999; 99DK-0001340.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Lehmann M;
XX
DR WPI; 2000-491161/43.
DR N-PSDB; AAA73232.
XX

1

```

XX SQ Sequence 467 AA;
Query Match 100.0%; Score 2470; DB 21; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.5e-240;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCPEPESHLMGWQYSPFFSLADE 60
Db 1 mgvfvlisiatlfgstgctalgrgnshscdtvdggycpeishlmwgtyspffslade 60
Qy 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 saispdvpkgrvtfvqvlshrgharyptsskkskysalialieaialknatafkgyaflkty 120
Qy 121 YTLGADDLTPGEOQMVNSGKIFVRRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltpgegqmvnsgikfrrykalarkivpfvrasgsdrviasaekfieqfsga 180
Qy 181 KLADPGANPHOASPVINVIIEGAGYNNNTLDHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
Db 181 kladpganphqaspviniiegagynntldhglctafeeselgddveantavfappir 240
Qy 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPCDLFTHDEWIQYDYLQSLG 300
Db 241 arleahlpgvnltdedvvnldmcpfdvtartsdatsqlspcdlftdhewiqdyqlsgl 300
Qy 301 KYYGYGAGNPLGPAQGVGVNELLARLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 kyygygagnglpagqgvgnellarlthspvqdhstnnhtldsnpatfplnatlyadfs 360
Qy 361 HDNTMVSIFFFALGYNLTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 hdntmvsiffalglyngtkplsttsvesieetdgyaaswtvpfaarayvemmqceakep 420
Qy 421 LVRVLVNDVRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWECFA 467
Db 421 lvrvlvndrvvplhcgvdkgkrdrdfveglsfarsggnwecfa 467

RESULT 3
AAY43170
ID AAY43170 standard; Protein; 467 AA.
XX AC AAY43170;
XX DT 06-JAN-2000 (first entry)
XX DE Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.
XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
XX KW consensus sequence.
XX OS Synthetic.
XX PN WO9948380-A1.
XX PD 30-SEP-1999.
XX PF 22-MAR-1999; 99WO-DK00154.
XX PR 23-MAR-1998; 98DK-0000407.
XX PR 19-JUN-1998; 98DK-0000806.
XX PR 18-SEP-1998; 98DK-0001176.
XX PR 22-JAN-1999; 99DK-0000091.
XX PR 22-JAN-1999; 99DK-0000093.
XX (NOVO ) NOVO-NORDISK AS.
XX PA Petersen S;
XX PI WPI; 1999-591030/50.
XX DR

```

```

DR N-PSDB; AAZ31521.
XX Preparing animal feed using a thermostable phytase
XX Example 3; Fig 10; 71pp; English.
XX CC This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
XX CC The invention relates to a process for preparing animal feed by
XX CC agglomerating feed ingredients with a thermostable phytase, which is
XX CC added before or during agglomeration. The thermostable phytase is useful
XX CC for expression in transgenic plants. These plants are useful in the
XX CC preparation of animal feed itself. The thermostable phytase allows animal
XX CC feed to be produced more efficiently, in addition to improved
XX CC phytase-expressing transgenic plants. These plants provide a feed
XX CC ingredient and a feed additive (phytase) simultaneously.
XX SQ Sequence 467 AA;

Query Match 99.1%; Score 2448; DB 20; Length 467;
Best Local Similarity 98.9%; Pred. No. 5.9e-238;
Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCPEPESHLMGWQYSPFFSLADE 60
Db 1 mgvfvlisiatlfgstgctalgrgnshscdtvdggycpeishlmwgtyspffslade 60
Qy 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 saispdvpkgrvtfvqvlshrgharyptsskkskysalialieaialknatafkgyaflkty 120
Qy 121 YTLGADDLTPGEOQMVNSGKIFVRRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltpgegqmvnsgikfrrykalarkivpfvrasgsdrviasaekfieqfsga 180
Qy 181 KLADPGANPHOASPVINVIIEGAGYNNNTLDHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
Db 181 kladpganphqaspviniiegagynntldhglctafeeselgddveantavfappir 240
Qy 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPCDLFTHDEWIQYDYLQSLG 300
Db 241 arleahlpgvnltdedvvnldmcpfdvtartsdatsqlspcdlftdhewiqdyqlsgl 300
Qy 301 KYYGYGAGNPLGPAQGVGVNELLARLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 kyygygagnglpagqgvgnellarlthspvqdhstnnhtldsnpatfplnatlyadfs 360
Qy 361 HDNTMVSIFFFALGYNLTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 hdntmvsiffalglyngtkplsttsvesieetdgyaaswtvpfaarayvemmqceakep 420
Qy 421 LVRVLVNDVRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWECFA 467
Db 421 lvrvlvndrvvplhcgvdkgkrdrdfveglsfarsggnwecfa 467

RESULT 4
AAB20527
ID AAB20527 standard; Protein; 467 AA.
XX AC AAB20527;
XX DT 05-DEC-2000 (first entry)
XX DE Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.
XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
XX KW temperature stability; pH profile; temperature profile; reaction rate;
XX KW specific activity; substrate specificity; substrate cleavage pattern;
XX KW substrate binding; position specificity; phytate degradation rate;
XX KW food; feed; phytate; manure.
XX OS Synthetic.

```

XX WO200043503-A1.
PN 27-JUL-2000.
XX 21-JAN-2000; 2000WO-DK00025.
XX 22-JAN-1999; 99DK-0000092.
PR 21-SEP-1999; 99DK-0001340.
XX (NOVO) NOVO NORDISK AS.
XX Lehmann M;
PI WPI; 2000-491161/43.
XX N-PSDB; AAA73234.
DR Novel phytases with improved properties such as temperature stability,
PT pH stability and substrate specificity, for use in pharmaceuticals and
PT compound foods and feeds -
XX Claim 4; Fig 8a-c; 240pp; English.
XX The present invention describes improved phytases, preferably with
CC increased thermostability, and methods for producing them. The methods
CC can be used for producing phytases with improved properties e.g.
CC temperature stability, pH stability, pH profile, temperature profile,
CC specific activity, substrate specificity, substrate cleavage pattern,
CC substrate binding, position specificity, the velocity and level of
CC release of phosphate from corn, reaction rate, phytate degradation rate,
CC and end level of released phosphate. The phytases can be used to produce
CC pharmaceutical compositions or compound food or feeds. The feed can be
CC used to reduce levels of phytate in animal manure, by converting it
CC into lower inositol phosphates and/or inositol and inorganic phosphate.
CC The present sequence represents a phytase sequence from the present
XX invention.
XX Sequence 467 AA;
SQ

Query Match 99.1%; Score 2448; DB 21; Length 467;
Best Local Similarity 98.9%; Pred. No. 5.9e-238;
Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTGTALGPRGNSHSCDVTGCGYOCFPEISHLWGQYSPFFSLADE 60
DB 1 mgvfvlslatlfgstgtalgrnshscdvtggyqcfpeishlwtgyspffslade 60

QY 61 SAISPDPVPGCRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAPFGKYAFILKTYN 120
DB 61 saispdpvgcrvtfvqvlshrhgaryptssaskaysalieaiqknatafkgkyafikty 120

QY 121 YTLGADDLTPFGQOMVNSGIFRYRKALARKIVFVRASGSDRVIASAEKIEGFQSA 180
DB 121 ytlgaddltpfgeqmvnsgikfrryrykalarxivfirasgsdrviasaekfiegfsa 180

QY 181 KLADPGANPHQASPVINVIPEAGVNTLIDHGLCTAFEESELGDDVEANFTAVFAPP 240
DB 181 kladpganphqaspvinviipeagvnyntldhglctafeeselgddveanftavfapp 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPEDTAVRTSDATQLSPFCDLTFHDEWIDYQLQSLG 300
DB 241 arleahlpvnltdedvnlmdmcpdvtavrtsdattqlspcdlftchdewidyqlqslg 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADES 360
DB 301 kygygagnplgpaqgvgnfneliarlthspvqgdhstnhtltdsnpatplnatlyads 360

QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqaakep 420

QY 421 LVRVLYNDRVPLHGGCVKGLGRCKRDDFVEGLSFARSGGNWEECFA 467

Db 421 lvrlyndrvplhggcvkglgrckrddfveglsfarsgnwecfa 467

RESULT 5
AAV69569
ID AAV69569 standard; Protein; 467 AA.
XX
AC AAV69569;
XX 19-APR-2000 (first entry)
XX Mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.
DE
XX
XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
KW thermostable; animal feed; monogastric animal; phytate phosphorus;
KW phosphate availability; consensus; mutant; mutein.
XX
XX Aspergillus terreus 9A1.
OS Aspergillus terreus cbs16.46.
OS Aspergillus niger var. awamori.
OS Aspergillus fumigatus ATCC13073.
OS Aspergillus fumigatus ATCC32722.
OS Aspergillus fumigatus ATCC58128.
OS Aspergillus fumigatus ATCC26906.
OS Aspergillus fumigatus ATCC32239.
OS Emericella nidulans.
OS Talaromyces thermophilus ATCC20186.
OS Myceliophthora thermophila.
OS Paxillus involutus NN005693.
OS Trametes pubescens NN9343.
OS Agrocybe pediades NN009289.
OS Peniophora lycii NN006113.
OS Thermomyces lanuginosa.
OS Synthetic.
XX
XX
XX Key Location/Qualifiers
FH 1..26
FT Peptide /note= "Phytase signal peptide from Aspergillus terreus
FT /note= cbs16.46"
FT Protein 27..467
FT /note= "Mature phytase-10-thermo[3]-Q50T-K91A"
XX
XX EP969089-A1.
XX 05-JAN-2000.
XX 23-JUN-1999; 99EP-0111949.
XX 29-JUN-1998; 98EP-0111960.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Brugger R, Lehmann M, Wyss M;
XX WPI; 2000-099429/09.
XX New stabilised enzyme formulation, useful for feed compositions for
PT monogastric animals -
XX
XX Example 5; Fig 19; 101pp; English.
XX The invention relates to a novel stabilised dry or liquid enzyme
CC formulation, comprising phytase (myo-inositol hexakisphosphate
CC phosphohydrolase) and one or more stabilising agents including
CC xylicol or ribitol; polyethylene glycols with a molecular weight of 600
CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
CC The stabilised phytase formulation is used in a method for preparing a
CC feed composition for monogastric animals (e.g., pigs, poultry) and
CC provides a monogastric animal with its dietary requirements of
CC phosphorus. Although a large amount of phosphate is present in animal

CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-10
 CC consensus sequence, phytase-10-thermo[31]-Q50T-K91A, which has a
 CC temperature optimum and melting point 4 degrees Celsius higher than that
 CC of phytase-10 (AAY69566). Its specific activity with phytate as a
 CC substrate is also strongly increased.

XX Sequence 467 AA;

Query Match 99.1%; Score 2448; DB 21; Length 467;
 Best Local Similarity 98.9%; Pred. No. 5.9e-238;
 Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCPEISHLMGQYSPFFSLADE 60
 DB 1 mgvfvllsialtfgstgaltgrgnschscdtvddgycpeshlwtgyspffslade 60
 QY 61 SAISPDPVPGKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGKYAFLKTYN 120
 DB 61 saispdpvpgkgrvtfvqvlshrgharyptsskksalsaliedalqknatafkgyaflkty 120
 QY 121 YTLGADDLTPFGQQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrrykalkarkivpfiragsdrviasaekfiqfgsa 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESLGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvilpeagayntldhglctafeeslgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNVLMDCPFDTVARTSDATQLSPFCDLTFHDEWTOYDYLQSLG 300
 DB 241 arleahlpgvnltedvnlmdcpfdtvtartsdatlspfcldtfhdewigydlqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnpplgpaqgvgfvnelliartlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyasaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLVNDRVVPLHGGCGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlvndrvvplhgcvdklgcrdrdfveglfsfarsggnwecfa 467

RESULT 6

AAB20533

ID AAB20533 standard; Protein; 467 AA.

XX AAB20533;

AC AAB20533;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.

XX WO200043503-A1.
 PN 27-JUL-2000.
 PD 21-JAN-2000; 2000WO-DK00025.
 PF 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PA Lehmann M;
 PI WPI: 2000-491161/43.
 DR N-PSDB; AAN73292.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 PS Disclosure; Fig 24a-c; 240pp; English.
 XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, the velocity and level of
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX Sequence 467 AA;

Query Match 98.9%; Score 2442; DB 21; Length 467;
 Best Local Similarity 98.7%; Pred. No. 2.4e-237;

Matches 461; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCPEISHLMGQYSPFFSLADE 60
 DB 1 mgvfvllsialtfgstgaltgrgnschscdtvddgycpeshlwtgyspffslade 60
 QY 61 SAISPDPVPGKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGKYAFLKTYN 120
 DB 61 saispdpvpgkgrvtfvqvlshrgharyptsskksalsaliedalqknatafkgyaflkty 120
 QY 121 YTLGADDLTPFGQQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrrykalkarkivpfiragsdrviasaekfiqfgsa 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESLGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvilpeagayntldhglctafeeslgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNVLMDCPFDTVARTSDATQLSPFCDLTFHDEWTOYDYLQSLG 300
 DB 241 arleahlpgvnltedvnlmdcpfdtvtartsdatlspfcldtfhdewigydlqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnpplgpaqgvgfvnelliartlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyasaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLVNDRVVPLHGGCGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467

Db 421 lvrvlndrvrplhgcavdklgrckrdffveglstfarsgnweecfa 467
 |||

RESULT 7
 AAB20534
 ID AAB20534 standard; Protein; 467 AA.

XX AC AAB20534;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX OS Synthetic.

XX PN WO200043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX DR WPI; 2000-491161/43.

DR N-PSDB; AAA73293.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Disclosure; Fig 25a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 98.6%; Score 2436; DB 21; Length 467;
 Best Local Similarity 98.5%; Pred. No. 9.5e-237;
 Matches 460; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGVFVLLSIALFGSTALGRGNSHSCDVTGQYOCFPEISHLWQYSPFFSLADE 60

Db 1 mgvfvlslatlfsgtsalgrgnshtcdvdygcfeishlwtsyffslade 60
 |||

QY 61 SALSPPDPKGRVTFVQVLSRHGARYPTSSKKYSALIEAQKNATAPKGYAFUKTYN 120

Db 61 salspdpkgrvtfvqvlshrgharyptsskaysallsaieaqlknatarkgyafuktyn 120
 |||

QY 121 YTLGADDLTPFGEQQVMNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIIEGQSA 180
 |||
 Db 121 ytlgaddltpfgeqmqvmsgkfyryrkalarakiyprfirasgdrviasaekfiegqsa 180
 |||
 QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTATFEESLGGDDVEANFTAVFAPPIR 240
 |||
 Db 181 kladpganphgaspvinviiepegagynntldhglctateestlgddveanftavfappir 240
 |||
 QY 241 ARLEAHLPGVNLTDDEVDVNLMDMCPDFTVARTSDATQLSPPCDLFTHDEWIOYDYLSLGL 300
 |||
 Db 241 arleahlpgvnltdedvvnldmmdmcpdftvartsdatsqlspcdlftdhewiodylslgl 300
 |||
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 |||
 Db 301 kyygygagnpqagvgvfneliarlthspvgdhtstnhtldsnpatfplnatlyadfs 360
 |||
 QY 361 HNTMVSITFFALGLNGKPLSTTSVESTIEEDGVAASWTVPFAARAYVEMMQCAEKPEP 420
 |||
 Db 361 hntmvsiffalglyngtkplsttsvestieedgvaswtvpfaarayvemmqceakep 420
 |||
 QY 421 LVRVLVNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 |||
 Db 421 lvrvlndrvrplhgcavdklgrckrdffveglstfarsgnweecfa 467
 |||

RESULT 8

AAB20531

ID AAB20531 standard; Protein; 467 AA.

XX AC AAB20531;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX OS Synthetic.

XX PN WO200043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX DR WPI; 2000-491161/43.

DR N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce

CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.3%; Score 2353; DB 21; Length 467;
 Best Local Similarity 94.9%; Pred. No. 2.2e-228;
 Matches 443; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTSGTALGRGNSHSCDVTGQYQCQFPEISHLWGOYSPFFSLADE 60
 Db 1 mgvfvlslatlfgstsgtalgrgnschscdtvggyqcfpeishlwgtyspyfslade 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNAFAFKGYAFLKTYN 120
 Db 61 saispdvpgkcrvtfvqlsrhgaryptskskaysaliaeqknatafkgyaflkty 120

QY 121 YTLGADDLTPFGEQMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKFTGFSQA 180
 Db 121 ytlgaddltpfgenqmvnsqikfrryrykalarkivpfiragsdrviasaeekfiegfsa 180

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpgsqphqaspvinviiepgsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPEDTARTSDATQLSFCDFTHDEWIQDYQLSLG 300
 Db 241 arleahlpvgvnltdedvvylnmdmcpdrtartsdatslfcalfthdewiqdyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFANIELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnpplgpaqvgfaneliarlthspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntwmisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqc 420

QY 421 LVRVLVNDRVVPLHGGVVDKLGRCRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 lvrvlvndrvvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

RESULT 9

AAB20532
 ID AAB20532 standard; Protein; 467 AA.

XX AC AAB20532;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX OS Synthetic.

XX PN W0200043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX XX

PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX DR WPI; 2000-491161/43.

XX DR N-PSDB; AAA73291.

XX PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX PS Disclosure; Fig 23a-c; 240pp; English.

XX CC The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX SQ Sequence 467 AA;

Query Match 95.0%; Score 2347; DB 21; Length 467;
 Best Local Similarity 94.8%; Pred. No. 8.9e-228;
 Matches 442; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTSGTALGRGNSHSCDVTGQYQCQFPEISHLWGOYSPFFSLADE 60
 Db 1 mgvfvlslatlfgstsgtalgrgnschscdtvggyqcfpeishlwgtyspyfslade 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNAFAFKGYAFLKTYN 120
 Db 61 saispdvpgkcrvtfvqlsrhgaryptskskaysaliaeqknatafkgyaflkty 120

QY 121 YTLGADDLTPFGEQMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKFTGFSQA 180
 Db 121 ytlgaddltpfgenqmvnsqikfrryrykalarkivpfiragsdrviasaeekfiegfsa 180

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpgsqphqaspvinviiepgsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPEDTARTSDATQLSFCDFTHDEWIQDYQLSLG 300
 Db 241 arleahlpvgvnltdedvvylnmdmcpdrtartsdatslfcalfthdewiqdyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFANIELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnpplgpaqvgfaneliarlthspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntwmisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqc 420

QY 421 LVRVLVNDRVVPLHGGVVDKLGRCRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 lvrvlvndrvvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

RESULT 10

AAB20523

ID AAB20523 standard; Protein; 441 AA.

XX AC AAB20523;

XX XX

05-DEC-2000 (first entry)

Consensus phytase 10 (Fcp10) SEQ ID NO:24.

Phytase; mutant; thermostability; mutation; mutagenesis; pH stability; temperature stability; pH profile; temperature profile; reaction rate; specific activity; substrate specificity; substrate cleavage pattern; substrate binding; position specificity; phytate degradation rate; food; feed; phytate; manure.

Synthetic.

WO200043503-A1.

27-JUL-2000.

21-JAN-2000; 2000WO-DK00025.

22-JAN-1999; 98DK-0000092.

21-SEP-1999; 98DK-0001340.

(NOVO) NOVO NORDISK AS.

Lehmann M;

WPI; 2000-491161/43.

Novel phytases with improved properties such as temperature stability, pH stability and substrate specificity, for use in pharmaceuticals and compound foods and feeds -

Example 2; Fig 4a-d; 240pp; English.

The present invention describes improved phytases, preferably with increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and end level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence represents a phytase sequence from the present invention.

Sequence 441 AA;

Query Match 94.9%; Score 2344; DB 21; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.6e-227;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 NNSCSDTVDGGYQCFPEISHLWQYSPFFSLADESAISPDVPRGCRVTFVQVLSRHGARY 86
Db 1 nshscdtdggyqcfpeishlwgyspffsladesaispdvprgcrvtfvqvlshrgary 60

QY 87 PTKSKKYSALIEATQKNATAFGKYAFLLKTYNTLGGADLLTPFGQQMWSGKIFRR 146
Db 61 ptksskksyalieatqknatafgkyafllktyntlgadlltpfgeqqmwsngikfrr 120

QY 147 YKALARKIVPFRASGSDRVASAEKPIEGFSQAKLADPGANPRQASPVINVIPEGAGY 206
Db 121 ykalarkivpfrasgsdrviasaekpiegfsqakladpganpqrqaspvinvipegag 180

QY 207 NNTLDHGLCTAFEESELGDDVEANFTAVFAPPTRARLEAHLPGVNLTDDEVNLMDCPF 266
Db 181 nntldhglctafeeselgddveanftavfapptrarleahlpvnltdedvnlmdmcpf 240

QY 267 DTVARTSDATLSPFCDFLTHDEWIOYDYQLSLGKYGYGAGNPLGPAQGVGFVNELIAR 326
Db 241 dtvartsdatlspfcdfldhewiodydyqlslgkygygagnplgpaqgvgfvnelliar 300

QY 327 LTHSPVDHTSTNHTLDGNPATFFPLNATLYADFSDNTMWSIFFALGLYNGTKPLSTTSV 386
Db 301 lthspvqdhstcnhldsnpatffplnatlyadfsdntmwsiffalglngtkplsttsv 360

QY 387 ESIEETDGYAASWTVPFAARAYVEMMOCEAEKEPLVRVLVNDRVVPLHGCGVDKLGRCR 446
Db 361 esieetdgyaaswtvpfaarayvemmoceaekeplvrvlvndrvvplhgcgvdklgrckr 420

QY 447 DDFVEGLSFARSGGNWEECF 467
Db 421 ddfveglsfarsggnweecfa 441

RESULT 11

AA43169

ID AA43169 standard; Protein; 467 AA.

XX

AC AA43169;

XX

DT 06-JAN-2000 (first entry)

XX

DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.

XX

KW Phytase; animal feed preparation; thermostable phytase; transgenic plant; consensus sequence.

OS Synthetic.

XX

PN WO9948380-A1.

XX

PD 30-SEP-1999.

XX

PF 22-MAR-1999; 99WO-DK00154.

XX

PR 23-MAR-1998; 98DK-0000407.

PR

PR 19-JUN-1998; 98DK-0000806.

PR

PR 18-SEP-1998; 98DK-0001176.

PR

PR 22-JAN-1999; 99DK-0000091.

PR

PR 22-JAN-1999; 99DK-0000093.

XX

PA (NOVO) NOVO-NORDISK AS.

XX

PI Petersen S;

XX

DR WPI; 1999-591030/50.

DR

DR N-PSDB; AA231520.

XX

PT Preparing animal feed using a thermostable phytase

XX

PS Example 3; Fig 9; 71pp; English.

XX

CC This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A. The invention relates to a process for preparing animal feed by agglomerating feed ingredients with a thermostable phytase, which is added before or during agglomeration. The thermostable phytase is useful for expression in transgenic plants. These plants are useful in the preparation of animal feed itself. The thermostable phytase allows animal feed to be produced more efficiently, in addition to improved phytase-expressing transgenic plants. These plants provide a feed ingredient and a feed additive (phytase) simultaneously.

XX

Sequence 467 AA;

Query Match 94.8%; Score 2342; DB 20; Length 467;
Best Local Similarity 94.4%; Pred. No. 2.8e-227;
Matches 441; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTGATLGRNSHSCDNDVGGYQCFPEISHLWGOYSPFFSLADE 60
Db 1 mgvfvlslatlfgstgatlgrnshscdndvggycfpeishlwgtyspysfslade 60

QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGKYAFELKTYN 120
 Db 61 saispdpvddcrvtvqvlshrgharyptssaskaysalieaiqknatafkgyafelktyN 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfiragsdrviasaekfiegfsga 180
 QY 181 KLADPGANPHOASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANTVAFAPPIR 240
 Db 181 kladpgsghaspsvinviiepgsgynnltldhgtctafedseigddveanftalfapair 240
 QY 241 ARLEAHLPGVNLTDVNNLMDMCPFDVTARTSDATLSPCDFLTHDEWIQDYLOSLG 300
 Db 241 arleadlpgvnltdedvnylmdmcpfdvtartsdatselspfcalfthdewiqdyloqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpaqgvgfaneliarlthspvgdhtstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdnmtisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlvndrvplhgcavdklgrckrdrdfveglsfarsggnwaecfa 467

RESULT 12

AAB20526
 ID AAB20526 standard; Protein; 467 AA.

XX AAB20526;
 AC AAB20526;

XX 05-DEC-2000 (first entry)
 XX Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,

CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 94.8%; Score 2342; DB 21; Length 467;
 Best Local Similarity 94.4%; Pred. No. 2.8e-227;

Matches 441; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 MGVFVWLLSIATLPGSTSGTALGPRGNSHSCDVTDDGGYQCPEISHLMGQYSPFFSLADE 60
 Db 1 mgvfvwllsialtlfgstsgtalgrgnshscdvtddggycpfeishlmwgtyspyfslade 60
 QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGKYAFELKTYN 120
 Db 61 saispdpvddcrvtvqvlshrgharyptssaskaysalieaiqknatafkgyafelktyN 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfiragsdrviasaekfiegfsga 180
 QY 181 KLADPGANPHOASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANTVAFAPPIR 240
 Db 181 kladpgsghaspsvinviiepgsgynnltldhgtctafedseigddveanftalfapair 240
 QY 241 ARLEAHLPGVNLTDVNNLMDMCPFDVTARTSDATLSPCDFLTHDEWIQDYLOSLG 300
 Db 241 arleadlpgvnltdedvnylmdmcpfdvtartsdatselspfcalfthdewiqdyloqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpaqgvgfaneliarlthspvgdhtstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdnmtisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlvndrvplhgcavdklgrckrdrdfveglsfarsggnwaecfa 467

RESULT 13

AAY69568

ID AAY69568 standard; protein; 467 AA.

XX AAY69568;
 AC AAY69568;

XX 19-APR-2000 (first entry)

XX Mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutin.

XX Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS Aspergillus niger var. awamori.

OS Aspergillus niger T213.

OS Aspergillus niger str. NRRL3135.

OS Aspergillus fumigatus ATCC13073.

OS Aspergillus fumigatus ATCC32722.

OS Aspergillus fumigatus ATCC58128.

OS Aspergillus fumigatus ATCC26906.

121	QY	121	YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFRASGRDVTASAEKFTIEGFQSA	181
121	Db	121	YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFRASGRDVTASAEKFTIEGFQSA	180
181	QY	181	KLADPGANPHQASPVINVIIEGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPP	240
181	Db	181	KLADPGSQHASPVINVIIEGAGYNNLTLDHGLCTAFEDSELGDDVEANFTALFAPAI	240
241	QY	241	ARLEAHLPGVNLTDDEVVNLMDMCPEDTVAFTSDATOLSPDCDLFTHDEWTOYDYLSQLG	300
241	Db	241	ARLEADIPGVLTDEDVVYLMDMCPEDTVAFTSDATOLSPDCDLFTHDEWTOYDYLSQLG	300
301	QY	301	KYGYGAGNPLGPAQGVGFYNELIARLTHSPQDHTSTNHTLDSNPATFPPLNATLYADFS	360
301	Db	301	KYGYGAGNPLGPAQGVGFANELIARLTHSPQDHTSTNHTLDSNPATFPPLNATLYADFS	360
361	QY	361	HNTWWSITFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVENMOC	420
361	Db	361	HNTMISIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVENMOCQAEK	420
421	QY	421	LVRVLVNDVRVPLHCGVDKLGKCRKDDFVEGLSTFARSGNNEECFA	467
421	Db	421	LVRVLVNDVRVPLHCGVAVDKLGKCRKDDFVEGLSTFARSGNNAECFA	467
RESULT 14				
AAW93380				
ID		AAW93380 standard; Protein; 467 AA.		
AC		AAW93380;		
XX		(first entry)		
DT		11-JUN-1999		
DE		Fungal phytase protein consensus DNA.		
XX		Phytase; consensus: myo-inositol hexakisphosphate; hydrolysis; food;		
KW		feed additive; variant; muten; feed; pharmaceutical.		
OS		Fungl.		
OS		Synthetic.		
PN		EP897985-A2.		
PD		24-FEB-1999.		
XX		15-JUL-1998; 98EP-0113176.		
PF		24-JUL-1997; 97EP-0112688.		
PR		(HOFF) HOFFMANN LA ROCHE AG F.		
PA		Lehmann M;		
PI		WPI; 1999-134647/12.		
XX		N-PSDB; AAX23022.		
DR		Preparation of a consensus protein, especially a phytase - using		
PT		programs to compare evolutionary similarity of sequences		
XX		Claim 8; Fig 2; 30pp; English.		
PS		This invention describes a novel process for the preparation of a		
CC		consensus protein. The specific example given in the specification is		
CC		that of a fungal phytase (myo-inositol hexakisphosphate) which hydro		
CC		lytase to valuable feed additives, with a fully defined amino acid		
CC		sequence given in the specification, or variant or muten. The metho		
CC		useful for improving protein properties by altering their sequence.		
CC		consensus protein and muten are useful in food, feed or pharmaceuti		
CC		compositions. This sequence represents the consensus phytase protein		
CC		in the method of the invention.		
XX		Sequence 467 AA;		
SQ				

Query Match 93.6%; Score 2311; DB 20; Length 467;
 Best Local Similarity 93.1%; Pred. No. 3.8e-224;
 Matches 435; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGATGPRGNHSCDVTGQYQCFPEISHLWGOYSPFFSLADE 60
 DB 1 mgvfvlisiatlfgstgatalprgnshscdvtggyqcfpeishlwgqyspyfslade 60

QY 61 SAISPDVPKGRVTFVOVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKTYN 120
 DB 61 saispdvpddrcvtvqlszngaryptsskkskaysalieaigknatafkgyafkty 120

QY 121 YTLGADDLTPFGEQMVNSGKIFRYRKALARKIVPFVRASGSDRVIASAEKFTGFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfyrrikalarkivpfirasgsdrviasaekfiegfgsa 180

QY 181 KLADPGANPHQASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVEANFTAFAPAIR 240
 DB 181 kladpgsqhgaspvdiviiepegsgyntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDVNVNLMDCPEDTVAFTSDATOLSPCEDLFTHDEWTOYDYLQSLG 300
 DB 241 arleadlpgvntldedvvnldmcpedvtartsdatsolspcedlftdewtoydylqslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNTWYSIFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 hdnwmisifalglngtckplsttsvesietdgyaaswtvpfagarayvemmqcaekep 420

QY 421 LVRVLVNDVRVPLHGGCVDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlvndrvvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

RESULT 15
 AAB20515
 ID AAB20515 standard; Protein; 467 AA.
 AC AAB20515;
 XX
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Consensus phytase SEQ ID NO:16.
 XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX
 OS Synthetic.
 XX
 PN WO200043503-A1.
 XX
 PD 27-JUL-2000.
 XX
 PF 21-JAN-2000; 2000WO-DK00025.
 XX
 PR 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Lehmann M;
 XX
 DR WPI; 2000-491161/43.
 DR N-PSDB; AAA73231.
 XX
 PT Novel phytases with improved properties such as temperature stability,

PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX Example 9; Fig 2a-c; 240pp; English.
 XX
 CC The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 93.6%; Score 2311; DB 21; Length 467;
 Best Local Similarity 93.1%; Pred. No. 3.8e-224;
 Matches 435; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGATGPRGNHSCDVTGQYQCFPEISHLWGOYSPFFSLADE 60
 DB 1 mgvfvlisiatlfgstgatalprgnshscdvtggyqcfpeishlwgqyspyfslade 60

QY 61 SAISPDVPKGRVTFVOVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKTYN 120
 DB 61 saispdvpddrcvtvqlszngaryptsskkskaysalieaigknatafkgyafkty 120

QY 121 YTLGADDLTPFGEQMVNSGKIFRYRKALARKIVPFVRASGSDRVIASAEKFTGFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfyrrikalarkivpfirasgsdrviasaekfiegfgsa 180

QY 181 KLADPGANPHQASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVEANFTAFAPAIR 240
 DB 181 kladpgsqhgaspvdiviiepegsgyntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDVNVNLMDCPEDTVAFTSDATOLSPCEDLFTHDEWTOYDYLQSLG 300
 DB 241 arleadlpgvntldedvvnldmcpedvtartsdatsolspcedlftdewtoydylqslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNTWYSIFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 hdnwmisifalglngtckplsttsvesietdgyaaswtvpfagarayvemmqcaekep 420

QY 421 LVRVLVNDVRVPLHGGCVDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlvndrvvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

Search completed: October 26, 2001, 16:40:03
 Job time: 4954 sec

us-09-488-265-26.rag

Sat Oct 27 15:25:38 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: October 26, 2001, 14:11:09 ; Search time 75.85 seconds
(without alignments)
814.587 Million cell updates/sec

Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGVFVLLSIATLFGSTSGT.....DFVEGLSFARSGNWECEFA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_16:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_unclassified:*
 - 13: sp_vertebrate:*
 - 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1902	77.0	465	3	O00092	O00092 aspergillus
2	1861	75.3	467	3	Q9UJZ7	Q9UJZ7 aspergillus
3	1855	75.1	467	3	Q93838	Q93838 aspergillus
4	1842	74.6	467	3	Q9HEQ0	Q9HEQ0 aspergillus
5	1826	73.9	466	3	O00100	O00100 aspergillus
6	1756	71.1	466	3	O00085	O00085 aspergillus
7	1719	69.6	466	3	O00096	O00096 talaromyces
8	1346.5	54.5	487	3	O00107	O00107 thielavia h
9	343.5	13.9	469	3	Q9Y846	Q9Y846 kluyveromyc
10	333.5	13.5	442	3	Q74677	Q74677 pichia anqu
11	301	12.2	463	3	O60172	O60172 schizosacch
12	212.5	8.6	460	5	Q9VV72	Q9VV72 drosophila
13	204	8.3	451	11	Q35217	Q35217 rattus norv
14	204	8.3	467	5	O96421	O96421 drosophila
15	196	7.9	481	11	Q92216	Q92216 mus musculu
16	195	7.9	487	4	O95172	O95172 homo sapien
17	194	7.9	487	4	Q9UNW1	Q9UNW1 homo sapien
18	192	7.8	487	4	Q9UGA3	Q9UGA3 homo sapien
19	179	7.2	449	13	Q92170	Q92170 gallus gall

20	176.5	7.1	453	5	O96420	O96420 drosophila
21	175.5	7.1	453	5	Q9W438	Q9W438 drosophila
22	167.5	6.8	198	3	Q9UTX1	Q9UTX1 schizosacch
23	154.5	6.3	468	10	O04509	O04509 arabidopsis
24	144	5.8	274	11	Q9JUD5	Q9JUD5 mus musculu
25	130	5.3	449	5	Q19076	Q19076 caenorhabdi
26	129.5	5.2	374	11	Q9JMG5	Q9JMG5 mus musculu
27	128.5	5.2	381	11	Q9QXG5	Q9QXG5 mus musculu
28	119.5	4.8	513	2	Q46334	Q46334 comamonas t
29	119	4.8	380	5	Q22525	Q22525 caenorhabdi
30	111.5	4.5	827	2	O24781	O24781 eubacterium
31	110.5	4.5	381	11	Q9QXH7	Q9QXH7 mus musculu
32	109	4.4	683	5	O00838	O00838 leishmania
33	109	4.4	707	5	O00839	O00839 leishmania
34	108.5	4.4	513	2	O24719	O24719 comamonas t
35	108	4.4	354	5	Q19709	Q19709 caenorhabdi
36	106.5	4.3	730	5	Q20826	Q20826 caenorhabdi
37	104	4.2	499	3	P87212	P87212 polyporacea
38	103.5	4.2	496	1	O34184	O34184 halobacteri
39	102.5	4.1	1013	14	O9ILX9	O9ILX9 retroperito
40	101.5	4.1	447	5	Q9U5U0	Q9U5U0 drosophila
41	101.5	4.1	447	5	Q9U5T8	Q9U5T8 drosophila
42	101.5	4.1	447	5	Q9TW19	Q9TW19 drosophila
43	101.5	4.1	602	2	Q9K5A0	Q9K5A0 erwinia chr
44	101	4.1	516	5	Q25327	Q25327 leishmania
45	101	4.1	763	2	Q9R6Y7	Q9R6Y7 anabaena sp

ALIGNMENTS

RESULT 1
O00092
ID O00092 PRELIMINARY: PRT: 465 AA.
AC O00092;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHVA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN=ATCC 34625;
RX MEDLINE=97288063; PubMed=9143104;
RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
RT phytase from the fungus Aspergillus fumigatus.";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMATIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC EMBL: U59804; AAB96872.1; -
CC HSSP: P34752; 1IHP.
CC InterPro: IPR000560; -
DR Pfam: PF00328; acid.phosphat. 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 26

```

FT CHAIN 27 465 3-PHYTASE A. 77.0%; Score 1902; DB 3; Length 465;
FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
FT SIMILARITY).
FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 359 359 PROTON DONOR (BY SIMILARITY).
FT DISULFID 30 39 BY SIMILARITY.
FT DISULFID 70 412 BY SIMILARITY.
FT DISULFID 213 463 BY SIMILARITY.
FT DISULFID 262 280 BY SIMILARITY.
FT DISULFID 434 442 BY SIMILARITY.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 50836 MW; 86FCID9058C9B2C9 CRC64;

Query Match 77.0%; Score 1902; DB 3; Length 465;
Best Local Similarity 77.1%; Pred. No. 2.5e-147;
Matches 360; Conservative 37; Mismatches 68; Indels 2; Gaps 2;

QY 1 MGVSFVLLSIATLFGSTGTALGPRGNHSCDTVDGGYOCFPEISHLWGOYSPFFSLADE 60
DB 1 MVTLTLLSAVLLSRVSAAPSSAG-SKSCDTVDLGYOCSPATSHLWGOYSPFFSLEDE 59
QY 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFELKTYN 120
DB 60 LSVSSKLKPCDRTILVQVLSRHGARYPTSSKSKYKLVTAIQANATDEKGFALFKTYN 119
QY 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
DB 120 YTLGADDLTPFGGQMVNSGKIFRYRYKALARSVPVFIKASGSDRVIAAEKFIQFQSA 179
QY 181 KLADPGANPHQASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 180 KLADPGAT-NRAAPASVVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 238
QY 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSPPCDFTHDEWIOYDYLQSLG 300
DB 239 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSPPCDFTHDEWIOYDYLQSLG 298
QY 301 KYGYGAGNPLGPAQGVGVNFVNLIELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
DB 299 KYGYGAGNPLGPAQGVGVNFVNLIELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 358
QY 361 HDNTMVSIFPAGLVNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP 420
DB 359 HDNTMVSIFPAGLVNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP 418
QY 421 LVRVLVNDVRVPLHGGCVGVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467
DB 419 LVRALINDVRVPLHGGCVGVDKLGCRKDDFVEGLSFARSGGNWEECEFS 465

RESULT 2
Q9U0Z7 PRELIMINARY; PRT; 467 AA.
AC Q9U0Z7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (EC 3.1.3.8).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RA Hongning W., Qi W., Jing X.;
RT "PCR, cloning and characterization of the phytase (phyA) gene of

```

```

RT Aspergillus niger (China Strain).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF218813; AAF25481.1; -.
DR HSSP: P34752; 1IHP.
DR InterPro: IPR000560; -.
DR Pfam: PF00328; acid_phosphat; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Signal; Lyase; Hydrolase.
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 75.3%; Score 1861; DB 3; Length 467;
Best Local Similarity 74.3%; Pred. No. 5.6e-144;
Matches 347; Conservative 47; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVSFVLLSIATLFGSTGTALGPRGNHSCDTVDGGYOCFPEISHLWGOYSPFFSLADE 60
DB 1 MGVSFVLLSIATLFGSTGTALGPRGNHSCDTVDGGYOCFPEISHLWGOYSPFFSLADE 60
QY 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFELKTYN 120
DB 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFELKTYN 120
QY 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
DB 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGANPHQASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSPPCDFTHDEWIOYDYLQSLG 300
DB 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSPPCDFTHDEWIOYDYLQSLG 300
QY 301 KYGYGAGNPLGPAQGVGVNFVNLIELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
DB 301 KYGYGAGNPLGPAQGVGVNFVNLIELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
QY 361 HDNTMVSIFPAGLVNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP 420
DB 361 HDNTMVSIFPAGLVNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP 420
QY 421 LVRVLVNDVRVPLHGGCVGVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467
DB 421 LVRVLVNDVRVPLHGGCVGVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467

RESULT 3
O93838 PRELIMINARY; PRT; 467 AA.
ID O93838
AC O93838
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PHYTASE.
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-57;
RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
RT "Phytase having high-affinity for phytic acid";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB022700; CAB19824.1; -.
DR HSSP: P34752; 1IHP.
DR InterPro: IPR000560; -.
DR Pfam: PF00328; acid_phosphat; 1.

```

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543EDC265C CRC64;

Query Match 75.1%; Score 1855; DB 3; Length 467;
 Best Local Similarity 73.4%; Pred. No. 1.7e-143;
 Matches 343; Conservative 51; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALPGRNHSCDVTDDGGYQCPEISHLWGQYSPFFSLADE 60
 DB 1 MGVSALLPLVLLSVTSGLAVPASRNOSTCDTVDGGYQCFSETSHLWGQYAPFFSLANK 60
 QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNAFAFKYAFKTYN 120
 DB 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNAFAFKYAFKTYN 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPVFRASGSDRVIASAEKFIQGFOSA 180
 DB 121 YSLGADDLTPFGEQELVNSGVKIFRYRYESLTRNIVPFRSSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGANPHQASPVINVLIPGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGGSSPKIDWVISEASTSNNTLDPGCTVFEDSELADDEANFTATVPSTR 240
 QY 241 ARLEAHLPGVNLTDVYVNLMDMCPFDVTARTSDATQLSPFCDLTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVSLTDEVTYVNLMDMCSFDITSTVDTKLSPFCDLTHDEWINYDYLQSLN 300
 QY 301 KYGYGAGNPLGPAAGVGVNELIARLTHSPVQDHTSNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATPLNATLYADFS 360
 QY 361 HDNTMVSIFALGNYGKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGNYGKPLSSTTAENITQDGFSSARTVPFAARMYVEMMOQSQEQEP 420
 QY 421 LVRVLNDRVVPLHCGVVDKLGRCRCKDDFVEGLSFARSGGNWEECFA 467
 DB 421 LVRVLNDRVVPLHCGPVDALGRCTRDSEFVKGLSFARSGGDWAECEFA 467

RESULT 4

ID Q9HEQ0 PRELIMINARY; PRT; 467 AA.
 AC Q9HEQ0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE PHYTASE.
 OS Aspergillus ficuum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L., An L., Wang Y., Yuan X.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY013315; AAG40885.1; -;
 SQ SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match 74.6%; Score 1842; DB 3; Length 467;
 Best Local Similarity 73.2%; Pred. No. 2e-142;
 Matches 342; Conservative 51; Mismatches 74; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALPGRNHSCDVTDDGGYQCPEISHLWGQYSPFFSLADE 60
 DB 1 MGVSALLPLVLLSVTSGLAVPASRNOSTCDTVDGGYQCFSETSHLWGQYAPFFSLANK 60
 QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNAFAFKYAFKTYN 120
 DB 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNAFAFKYAFKTYN 120

QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPVFRASGSDRVIASAEKFIQGFOSA 180
 DB 121 YSLGADDLTPFGEQELVNSGVKIFRYRYESLTRNIVPFRSSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGANPHQASPVINVLIPGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGGSSPKIDWVISEASTSNNTLDPGCTVFEDSELADDEANFTATVPSTR 240
 QY 241 ARLEAHLPGVNLTDVYVNLMDMCPFDVTARTSDATQLSPFCDLTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVSLTDEVTYVNLMDMCSFDITSTVDTKLSPFCDLTHDEWINYDYLQSLN 300
 QY 301 KYGYGAGNPLGPAAGVGVNELIARLTHSPVQDHTSNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATPLNATLYADFS 360
 QY 361 HDNTMVSIFALGNYGKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGNYGKPLSSTTAENITQDGFSSARTVPFAARMYVEMMOQSQEQEP 420
 QY 421 LVRVLNDRVVPLHCGVVDKLGRCRCKDDFVEGLSFARSGGNWEECFA 467
 DB 421 LVRVLNDRVVPLHCGPVDALGRCTRDSEFVKGLSFARSGGDWAECEFA 467

RESULT 5

ID Q00100 PRELIMINARY; PRT; 466 AA.
 AC Q00100;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE 3-PHYTASE PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE).
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 116.46;
 RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B.,
 RA Broger C., van Loon A.P.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U60412; AAB58465.1; -;
 DR HSSP; P34752; IHP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 466 3-PHYTASE.
 FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY
 FT ACT_SITE 83 83 SIMILARITY).
 FT ACT_SITE 361 361 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT DISULFID 31 40 BY SIMILARITY.
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	207	207	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	376	376	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	466 AA;	51055 MW;	F2AECCE1AF7C22C4	CRC64;
Query Match					
Best Local Similarity 73.9%; Score 1826; DB 3; Length 466;					
Matches 339; Conservative 44; Mismatches 83; Indels 0; Gaps 0;					
QY	1	MGVFFVLLSIATLFGSTGTALGPRGNSHSDTVGGYQCFPEISHLWGOYSPFFSLADE	60		
DB	1	MGVFFVLLSIATLFGSTGTALGPRGNSHSDTVGGYQCFPELSHKWGLYAPYFSLODE	60		
QY	61	SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFKLTYN	120		
DB	61	SPFPLDVPDCHITFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFKLTYN	120		
QY	121	YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGDRVIAAEKFEQFQSA	180		
DB	121	YSGSENLPFGNQDLGCAQFYRRYDFTLRHINPFVRADSRVHESAEKFEQFQSA	180		
QY	181	KLADPGANPHQASPVNLIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR	240		
DB	181	ROGDPHANPHQSPRVDDVVIPEGTAYNTLHSHCTAFESTVGDGDAVNFTAVFAPAI	240		
QY	241	ARLEAHLPGVNLTDDEVNLMDCPFTVARTSDATQLSPFCDFTHDEWIQYDYLQSLG	300		
DB	241	KRLAADLPGVQLSADDVNLNLMDCPFTVARTSDATQLSPFCDFTHDEWIQYDYLQSLG	300		
QY	301	KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	360		
DB	301	KYGYGGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	360		
QY	361	HDMTWSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP	420		
DB	361	HDSNLVSIFWALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP	420		
QY	421	LVRVLNDRVMPHLCGCKRDAFVAGLSFAQAGGNWADCF	466		
DB	421	LVRVLNDRVMPHLCGCKRDAFVAGLSFAQAGGNWADCF	466		
RESULT 6					
ID	000085	PRELIMINARY; PRT; 466 AA.			
AC	000085;				
DT	01-JUL-1997 (T-EMBLrel. 04, Created)				
DT	01-JUL-1997 (T-EMBLrel. 04, Last sequence update)				
DT	01-MAY-2000 (T-EMBLrel. 13, Last annotation update)				
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-				
DE	PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE				
DE	PHOSPHOHYDROLASE A).				
GN	PHYA.				
OS	Aspergillus terreus.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=33178;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=9A1;				
RX	MEDLINE=97117792; PubMed=9025298;				
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;				
RT	"the phytase subfamily of histidine acid phosphatases: isolation of				
RT	genes for two novel phytases from the fungi Aspergillus terreus and				
RT	Myceliophthora thermophila."				
RL	Microbiology 143:245-252(1997).				
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE				
CC	FROM PHYTATE (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-				
CC	INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.				

CC	-1- SUBCELLULAR LOCATION: SECRETED.				
CC	-1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO 7.5, WITH MAXIMAL ACTIVITY AT PH 5.5. ALSO ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO MORE ACIDIC PH VALUES.				
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.				
CC	EMBL; U59805; AA852507.1; -.				
DR	HSSP; P34752; 1IHP.				
DR	InterPro; IPR000560; -.				
DR	Pfam; PF00328; acid.phosphat.1.				
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.				
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.				
KW	Hydrolase; Glycoprotein; Signal.				
FT	SIGNAL 1 19 POTENTIAL.				
FT	CHAIN 20 466 3-PHYTASE A.				
FT	ACT_SITE 82 REQUIRED FOR BINDING SUBSTRATE (BY SIMILARITY).				
FT	ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).				
FT	ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).				
FT	DISULFID 31 40 BY SIMILARITY.				
FT	DISULFID 71 414 BY SIMILARITY.				
FT	DISULFID 215 465 BY SIMILARITY.				
FT	DISULFID 264 282 BY SIMILARITY.				
FT	DISULFID 436 444 BY SIMILARITY.				
FT	CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).				
SQ	SEQUENCE 466 AA; 51093 MW; 21CDCB559C96AE66 CRC64;				
Query Match 71.1%; Score 1756; DB 3; Length 466;					
Best Local Similarity 69.7%; Pred. No. 2.2e-135;					
Matches 325; Conservative 54; Mismatches 87; Indels 0; Gaps 0;					
QY	1	MGVFFVLLSIATLFGSTGTALGPRGNSHSDTVGGYQCFPEISHLWGOYSPFFSLADE	60		
DB	1	MGFLAIVLSVALLFRSTSGTLPGRGKHSNCNSVDHGQCFPELSHKWGLYAPYFSLODE	60		
QY	61	SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFKLTYN	1		
DB	61	SPFPLDVPDCHITFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFLOSYN	1		
QY	121	YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGDRVIAAEKFEQFQSA	1		
DB	121	YSLDSEELTPFGNQDLGCAQFYRRYNALTRHINPFVRATDSRVHESAEKFEQFQTA	1		
QY	181	KLADPGANPHQASPVNLIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR	2		
DB	181	RODDHANPHQSPRVDDVVIPEGSAYNNTLHSHCTAFESTVGDGDAVNFTAVFAPAI	2		
QY	241	ARLEAHLPGVNLTDDEVNLMDCPFTVARTSDATQLSPFCDFTHDEWIQYDYLQSLG	3		
DB	241	QRLAADLPGVQLSTDDVNLNLMDCPFTVARTSDATQLSPFCDFTHDEWIQYDYLQSLG	3		
QY	301	KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	3		
DB	301	KYGYGGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	3		
QY	361	HDMTWSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEP	4		
DB	361	HDSNLVSIFWALGLYNGTKPLSTTSVESISQTDGYAAWTVFPAARAYVEMMOCEAKEP	4		
QY	421	LVRVLNDRVMPHLCGCKRDAFVAGLSFAQAGGNWADCF	466		
DB	421	LVRVLNDRVMPHLCGCKRDAFVAGLSFAQAGGNWADCF	466		
RESULT 7					
ID 000096 PRELIMINARY; PRT; 466 AA.					

000096;
 01-JUL-1997 (TReMBLrel. 04, Created)
 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 PHA).
 OS Talaromyces thermophilus.
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutiales; Trichocomaceae; Talaromyces.
 OX NCBI_TaxID=28565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20186;
 RX MEDLINE=98007872; PubMed=9349716;
 RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,
 van Loon A.P.G.M.;
 RA "Cloning of the phytases from *Emericella nidulans* and the thermophilic
 fungus *Talaromyces thermophilus*."
 RL Biochim. Biophys. Acta 1353:217-223(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTASE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- MISCELLANEOUS: HAS POTENTIAL AS FOOD AND FEED ADDITIVE. IT CAN
 FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
 USED AS FOOD FOR MONOGASTRIC ANIMALS.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59802; AAB96873.1; -;
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; -;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal
 FT SIGNAL 1 14
 FT CHAIN 15 466
 FT ACT_SITE 79 79
 FT ACT_SITE 357 357
 FT DISULFID 28 37
 FT DISULFID 68 410
 FT DISULFID 212 461
 FT DISULFID 261 278
 FT DISULFID 432 440
 FT CARBOHYD 204 204
 FT CARBOHYD 269 269
 FT CARBOHYD 335 335
 FT CARBOHYD 348 348
 FT CARBOHYD 372 372
 SQ SEQUENCE 466 AA; 51450 MW; FC4575B521A5C929 CRC64;
 Query Match 69.6%; Score 1719; DB 3; Length 466;
 Best Local Similarity 69.3%; Pred. No. 2.3e-132;
 Matches 325; Conservative 45; Mismatches 91; Indels 8; Gaps 3;
 QY 1 MGFWVLLS--IATLFGSTGCTALGRGNSHSDTVDGGYQCPPEISHLWGQSPFSLA 58
 DB 1 MSLLLVLSGLVALYYSRN-----PHVDSHSCNTVEGGYQCPPEISHSGWQSPFSLA 55
 QY 59 DESAISPDPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAPKGYAFLEKT 118
 DB 56 DQSEISDPDPQNKITFVQLLSRHGARYPTSSKSKYSALIEAIQKNATAPKGYAFLEKT 115
 QY 119 YNVTGLADDTLPGEQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAKFIEGFQ 178
 DB 116 YRQLGANDLTPEGNQMIQLGKFNHYKSLARNAPFVRCSGSDRVIASGRLFIEGFQ 175
 QY 179 SAKLADPGANPHQASPVINVIIEGAGYNNLTDLHGCLTAPESEELGDDVEANFTAVFAPP 238
 DB 176 SAKVLDPHSKDHPPTINVIIEEGSYNNLTGSCPFVEDSSGGHDGAEKAFQFAPA 235

QY 239 IRARLEAHLPGVNLTDQVNVNMDMCPDFTVARTSDATQLSFPDCLFTHDEWIDYDLOS 298
 DB 236 ILEKIRKHLPGVDLAVSDVPYLMDCPPETLARNHTDT-LSPFCALSTQEEWQAYDYS 294
 QY 299 LKYYGYGAGNPLGPAQGVGFVNELIARTHSPVODHTSTNHTLDSNPATFPLNATLYAD 358
 DB 295 LGKYYGNGGNGPLGPAQGVGFVNELIARTHSPVQDYTTVNTLDSNPATFPLNATLYAD 354
 QY 359 FSHDNTMVSIFALGLYNGTKPLSTSVESIEETDGYAASWTVPFAARAYVEMMOCEAK 418
 DB 355 FSHDNTMVSIFAALGLYNGTKPLSTSVESIEETDGYAASWTVPFGGRAYIEMMQCDDSD 414
 QY 419 EPLVRLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSNGNWECEFA 467
 DB 415 EPVRLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSNGNWECEFA 463
 RESULT 8
 ID 000107 PRELIMINARY; PRT; 487 AA.
 AC 000107;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 PHA).
 GN Thielavia heterothallica.
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Chaetomiaceae; Thielavia.
 OX NCBI_TaxID=78579;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97177792; PubMed=9025298;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,
 van Loon A.P.G.M.;
 RA "The phytase subfamily of histidine acid phosphatases: isolation of
 genes for two novel phytases from the fungi *Aspergillus terreus* and
Myceliophthora thermophila."
 RL Microbiology 143:245-252(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTASE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO
 ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA
 SHIFTED TO MORE ACIDIC PH VALUES.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59806; AAB52508.1; -;
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; -;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 75
 FT CHAIN 1 487
 FT DOMAIN 267 270
 FT DOMAIN 423 433
 FT ACT_SITE 75 75
 FT ACT_SITE 76 76
 FT ACT_SITE 368 368
 FT DISULFID 26 35
 FT DISULFID 64 421
 FT DISULFID 208 485
 FT DISULFID 260 289
 FT DISULFID 456 464
 FT CARBOHYD 165 165
 FT CARBOHYD 200 200

```

RESULT 10
ID 074677 PRELIMINARY; PRS; 442 AA.
AC 074677;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE (EC 3.1.3.2).
GN PHO1.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AYCC34438;
RX MEDLINE=98386672; PubMed=9720203;
RA Phongdara A., Merckelbach A., Keup P., Gellissen G., Hollenberg C.P.;
RT "Cloning and characterization of the gene encoding a repressible acid
RT phosphatase (PHO1) from the methylotrophic yeast Hansenula
RT polymorpha.";
RL Appl. Microbiol. Biotechnol. 50:77-84(1998).
DR EMBL; AF051161; AAC62537.1; -
DR HSSP; P34755; 1QF8.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.

```

SQ SEQUENCE 442 AA; 49370 MW; 7087D91A85B05C31 CRC64;

Query Match 13.5%; Score 333.5; DB 3; Length 442;
 Best Local Similarity 25.8%; Pred. No. 4.6e-19;
 Matches 106; Conservative 63; Mismatches 159; Indels 83; Gaps 16;

QY 66 DVPKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQK-NATAFKGYAFKLTNYNTLG 124
 DB 55 DTPPHCEIEQAQLEMRHGERFPTKSSGGKKFYDKLAKANITDYKGLAFIEDLEYFVP 114
 QY 125 ADDLTPFGEQWVNSGI---KF---YRRYKAL--ARKIVFVRASGSDRVIAAEKFI 174
 DB 115 DSDNYELETTRGLYSGLNAPKFGTYRLERYDSLVDTSVLPFIAAS-EDRVVDIARFSG 173
 QY 175 EGFQSAKLADPGANPHQASPVINVI---IPEGAGYNTLDHGLC---TAFERSELGDD 226
 DB 174 RGFGCPDYA-----TSCSIQVNETDTSGKANALITKDN--CPTVNSFFYDYSFGDE 223
 QY 227 V---EANTFVAPPFIRARLEAHLPGVNLTDDEVVNLMDMCPFDTVARTSDATQLSPFCD 283
 DB 224 IFQREAD-----RLNELSPGFNITADDITMGTYCAYETNYKGH-----SSPCD 267
 QY 284 LFTDHEWIOYDLOSLGKYGYGAGNPLGPAQGVGVFNELIARLTHSPVQDHTSTNHTLD 343
 DB 268 ALSREAFALQYNDVTKFYQFGPCYNMSAVAGGVYAN----- 305
 QY 344 SNPTFPL---NATLYADFSDNTMWSIFFALGYNGTKPLSTTSVESIETDGYAASWT 400
 DB 306 ---ATAKLQEDGKLWFSFSDNDLLNVTALGLITDYE-----LGTEDVDFHRFKTSEL 358
 QY 401 VPFARAYVEMMOCEAKEPLRVLVNDRVVPLHCGGVDKLGRCKRDDFVE 451
 DB 359 VPOGARLIIEKLNC--SDTSEVRTILNDKVVVPVPGCSGGPGYSCPLEDYLD 407

RESULT 11

ID 060172 PRELIMINARY; PRT; 463 AA.

AC 060172;
 DT 01-AUG-1998 (T-EMBLrel. 07, Created)
 DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
 DE PROBABLE ACID PHOSPHATASE (EC 3.1.3.2).
 GN SPC21H7.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; AL023286; CAA18863.1; -;
 DR HSSP; P34755; 1QFX.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein.
 FT ACT_SITE 69 69 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 340 340 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 463 AA; 52758 MW; 6C41AE422C6D624A CRC64;

Query Match 12.2%; Score 301; DB 3; Length 463;
 Best Local Similarity 24.3%; Pred. No. 2.3e-16;
 Matches 106; Conservative 71; Mismatches 198; Indels 52; Gaps 13;

QY 51 YSPFSLADESAISDPVPGCRVTFVQVLSRHGARYPTS--SKSKKYS-----L 98
 DB 43 HEPYFDGLDSA-----FPETCEIQOQHLLQRHGSRNPTGDTVATDVSQYLNNEFOEKL 97
 QY 99 IEAIQKNATAFKGYAFKLTNYNTL---CADDLTPFGEQWVNSGIKFYRRYKALKARKIV 155
 DB 98 NGSPVNSYFENPCFIKQWTPVIDAENADQLSRGRLEUFDLGRQLYQRYKLFDSYV 157
 QY 156 PFVRASGSDRVIAAEKFIIEGFSAKLADPGANPHQASPVINVIPEG--AGYNTLTDHG 213
 DB 158 YDINTAEQERVVESAKWFTYGLFGDKMYE-----KTNFILLSEKGAAGANSLSMYN 208
 QY 214 LCTAFEESEL-----GDDVEANFTAVFAPPFIRARLEAHL-PGVNLTDEDVNLMDMCPEDT 268
 DB 209 ACPVFKDNFNHKNATDAAHAVVRNFIIEPIVNLAKYFDSSYKLTINDVRSIFYICEVEI 268
 QY 269 VARTSDATQLSPFCDLFTDHEWIOYDLOSLGKYGYGAGNPLGPAQGVGVFNELIARLT 328
 DB 269 AIKDH-----SDFCSITFPSEFLNFEYDSDLDQAYGGGVSSEWASTLGGAYINNLA 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMWSIFFALGYNGTKPLSTTSVES 388
 DB 323 -----RNVTNPDEDRK-----VFLATHDSNIIPVEAALGFFPDITPQNPLPTDK 367
 QY 389 IEETDGYAASWTVPFAARAYVEMMOCEAKEPLRVLVNDRVVPLHCGGVDKLGR-----C 444
 DB 368 NIYTSOKTSSFSFVFNAGNLITELFFC-SDSKYYVRHLVNOQVYPLIDCGYSGPSTGSLC 426
 QY 445 KRDDFVEGLSEFSGGN 461
 DB 427 ELQAYLNSPIRANSTSN 443

RESULT 12

QYVW72

ID 09VW72 PRELIMINARY; PRT; 460 AA.

AC 09VW72;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
 DE MIPPI PROTEIN.
 GN MIPPI OR CG4123.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

```
QY      52 SPFFS-----LADESA--TSPDVPKG-CKVTVPVQLSVLRHARGARYPSSSKKYS 96  
      ||| : |  
Db     14 SPYFGTKRYEDVNPNLLGDPVAPRRDPELLAGCTPFVOLVALIRHTGYPTTKQIRKL 73  
      : |  
QY     97 ALTEAQKNATAEPKYAFCLKTYNLTGLADDLPFGEQQMVNSGIKFERYRALKARKI VP 156  
      ||| : |  
Db    74 QLOCLLOTRESVDGGSRVAALDQWPLUYDD---WMDDQLVEKGRDMQLAIRLAALFP 130  
      : |  
QY   157 --FVR-----ASGSDRVIASAEKFTIEGFOSAKLPADPGANPHQAQSFVINVLTPESGA 205
```

Db 131 DLFCRENYGRRLITSSKRCVDSAAFLQGLW--QHVHPGLPP---PDVSDMECDPPR 184
 QY 206 YNNITL---DHGLCAFEESELGDVVEANFTAVF-----APPRARLE-----AHLPG 249
 Db 185 VNDKLMRFEDH--CKFL-----TEVERNATLYHVEAFKTPGEMQTLKKVAATLQVPV 237
 QY 250 VNLTDVNNLMDMCPFDVTARTSDATOLSPFCDLTHDEWIQYDYLQSLGKYGYGAGN 309
 Db 238 NNLNADLIQVAFFTCGFDLAIQGVH---SPWCDVDFDADAKVLEYLNDLKOYWRSTGY 293
 QY 310 PLGPAQGVGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIF 369
 Db 294 AINSRSCNLFQDFLHDKHKAVEQKORSQ-----PVSSVILQFQHAETLLPLL 342
 QY 370 FALGLYNGTKPLSTSVESIEETGCAASWTVPFAARAYVEMOCEAEKEP-----LVRVL 425
 Db 343 SLMGVFKDKPLTAYNEEQVHRE-FRSGHIVPYASNLIFVLYHCEDAQTPOEKFQIOML 401
 QY 426 VNDRVVPL 433
 Db 402 LNEKVLP 409

RESULT 14
 O96421 ID O96421 PRELIMINARY; PRT; 467 AA.
 AC O96421;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 1.
 GN MIPPI OR CG4123.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046913; AAD02436.1;
 DR FlyBase; FBgn0026061; Mippl.
 DR InterPro; IPR000560; Mippl.
 DR Pfam; PF00328; acid_phosphat; 1.
 SQ SEQUENCE 467 AA; 53602 MW; 1A0571B9958B2632 CRC64;

Query Match 8.3%; Score 204; DB 5; Length 467;
 Best Local Similarity 22.9%; Pred. No. 2e-08;
 Matches 103; Conservative 66; Mismatches 173; Indels 108; Gaps 20;

QY 66 DVPK-----GCRVTFVQVLSRHGARYPTSS---KSKYSALIEAIQKNATAFKKGK-----112
 Db 46 DIDKQVLPFCQPKQMWIFRHGTRLPKPKSMINKASRVAELRDLIINNQQVARTKPTDA 105
 QY 113 -----YAFKTYNTLIGADLLPFGEQGMVNSGKIFRYRYKALARKIVPV-----158
 Db 106 LCQTDLIAIKLWKN---SSITPDMEEXLTAQGYEDLRGTAKLYQRYTTLTANTNDY 162
 QY 159 ---RAGSDRVIASAEKFTG-FQSAKLADPGANPHO-----ASPVINY-IIPEGA 204
 Db 163 YQFRHTDTORTTESKFAEGFLGSONAHPVEIPKQDLLLRPYDYCSSFKVNYKDEGS 222
 QY 205 GYNNITLDHGLCTAFEESELGDVVEANFTAVFAPPRARLEAHLPG--VNLTDVNNLMD 262
 Db 223 EYK-----FHQSCLYNDLAD-----ISTRL-----GFLYTLLEADIKLYMD 260
 QY 263 MCPFDVTARTSDATOLSPFCDLTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVFVNE 322
 Db 261 MCRYE---QAMVNDNRNVCWGAFLPEQITVFEYLEDLYKYGYGFFPENALNCRLVQD 317

QY 323 LIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKPLS 382
 Db 318 LLTHLS-NPVSPH-----VVXHFHSTGLLLTLLTALGKXKDDIKUR 357
 QY 383 TTSVESIETDGYAASWTVPFAAR--AYVEMOCEAEKEPLRVLVNDRVVPPLHGGGVDK 440
 Db 358 ADNYDSL-TSRWKSLLDPFAANEVAVKYDLPAFLDREKVV-FELNQOAVQLDWCVS--413
 QY 441 LGRCKRDRDVE-----GLSFARSGG 460
 Db 414 -GLCKWSDVLEKYKTIADADCGEYCYRTGG 442

RESULT 15
 O922L6 ID O922L6 PRELIMINARY; PRT; 481 AA.
 AC O922L6;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
 GN MINPPI OR MIPP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046908; AAD02434.1;
 DR MGD; MGI:1336159; Minppl.
 DR InterPro; IPR000560; Minppl.
 DR InterPro; IPR000886; Minppl.
 DR Pfam; PF00328; acid_phosphat; 2.
 DR PROSITE; PS00034; ER_TARGET; UNKNOWN_1.
 SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;

Query Match 7.9%; Score 196; DB 11; Length 481;
 Best Local Similarity 21.0%; Pred. No. 9.3e-08;
 Matches 100; Conservative 71; Mismatches 194; Indels 112; Gaps 20;

QY 11 ATLFGSTGTALPRGNHSCDVTVDGYQCFPEISHLWGOYSPFFSLAD--ESA-----I 63
 Db 21 AALLSSFARCSLPRGD-----PVASVL-----SPVFGTKTRYEDANPWL 62
 QY 64 SPDVPK-----GCRVTFVQVLSRHGARYPTSSSKYSALIEAIQKNATAFKGYAF 115
 Db 63 DPVAPRDPPELLAGTCTPVOLVALIRHGTTRYPTTKQIRKLKQLQGLLORESRDGSQA 122
 QY 116 LKTYNTLIGADLLPFGEQGMVNSGKIFRYRYKALARKIVP--FVR-----ASGSD 164
 Db 123 AALAEMPLWGD---WMDGOLVEKGRQDMQLALRALAALFPDLFSRENYDRRLITSSKH 179
 QY 165 RVTASAEKIEGFQSAKLADPGANPHOAS-----PVINVIPEGAGYNNITLDHGLCTAF 218
 Db 180 RCVDSAAFLQGLW---QHYHPLGLPPDPVSDMECGPPRINDKL-----MRFDDH--CEKF 229
 QY 219 ESELDGDDVVEANFTAVFAPPRARLEAHLPG-----VNLTDVNNL- 260
 Db 230 L-----TDVERNATALY-----HYEAFYTGEMQVKLVKVAATLQVPMNSLNADLIQVA 278
 QY 261 MDMCPFDVTARTSDATOLSPFCDLTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVFV 320
 Db 279 FFTCSFDLAIKGVH---SPWCDVDFDADARVLEYLNDLKOYWKRSYGYTINSRSCNLF 334
 QY 321 NELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKP 380
 Db 335 QDIFLHDKAVEQKORSQ-----PVSSPVIQFQHAETLLPLLISLMGYFKDKEP 383

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:27 ; Search time 48.78 Seconds
(without alignments)
682.417 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDVTGQCFETSHLW.....DFVEGLSFARSGGNWACFA 437

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1688	72.7	467	1 JN0889	3-phytase (EC 3.1.1.3.8)
2	1684	72.6	467	1 JN0856	3-phytase (EC 3.1.1.3.8)
3	1665	71.7	441	1 JN0482	3-phytase (EC 3.1.1.3.8)
4	356	15.3	467	2 S53476	acid phosphatase
5	356	15.3	467	2 S48996	acid phosphatase
6	353	15.2	467	2 S48996	acid phosphatase
7	347	15.0	467	1 PABYCC	acid phosphatase
8	337	14.5	468	2 S52495	acid phosphatase
9	334.5	14.4	468	2 JC4285	acid phosphatase
10	307.5	13.2	479	1 JN0890	acid phosphatase
11	303.5	13.1	479	1 JN0715	acid phosphatase
12	298	12.8	453	1 A25326	3-phytase (EC 3.1.1.3.8)
13	290	12.5	463	2 S14119	acid phosphatase
14	285	12.3	463	2 T39929	acid phosphatase
15	154	6.6	468	2 A86233	thiamin-repressibl
16	149.5	6.4	465	2 JEO369	hypothetical prote
17	125	5.4	465	2 JEO369	histidine acid pho
18	113.5	4.9	423	2 T19118	acid phosphatase h
19	111	4.8	423	2 A33395	acid phosphatase
20	111	4.8	693	1 S61067	hypothetical prote
21	108.5	4.7	421	2 S14742	homocarnitine hydr
22	107.5	4.6	381	2 JH0152	acid phosphatase
23	106	4.6	2205	2 T08615	acid phosphatase
24	105	4.5	537	2 S54770	aggregation factor
25	105	4.5	888	2 T46726	secreted acid phos
26	104.5	4.5	413	2 T18945	secreted acid phos
27	104.5	4.5	423	1 S06167	hypothetical prote
28	104	4.5	1025	2 G81722	acid phosphatase
29	102	4.4	1081	2 T51613	polymorphic membra
					DNA mismatch repai

30	101	4.4	419	2 T39920	probable glucanase
31	100	4.3	407	2 T40582	hypothetical prote
32	100	4.3	556	2 S67097	probable membrane
33	98.5	4.2	583	2 S46265	cysteine proteinase
34	98.5	4.2	1527	2 JE0336	canaliculus multia
35	97.5	4.2	336	2 A24430	glyceraldehyde-3-p
36	97	4.2	730	2 T16455	hypothetical prote
37	97	4.2	1458	2 A45665	adult-specific bru
38	96.5	4.2	844	2 T37690	hypothetical prote
39	96.5	4.2	1350	2 G36793	genome polyprotein
40	96.5	4.2	2290	1 GNNYE	hypothetical prote
41	95.5	4.1	357	2 T07851	ananain (EC 3.4.22
42	95.5	4.1	394	1 KIBSGM	phosphoglycerate k
43	95.5	4.1	647	2 A84265	hypothetical prote
44	95.5	4.1	986	2 T33135	hypothetical prote
45	95	4.1	455	2 B71480	probable agx-1 hom

ALIGNMENTS

RESULT 1

JN0889
3-phytase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
C:Species: Aspergillus awamori
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0889
R:Piddington, C.S.; Houston, C.S.; Palohelimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0889
A:Molecule type: DNA
A:Residues: 1-467 <PID>
A:Cross-references: GB:I02421; NID:q166518; PIDN:AAAL6898.1; PID:q166519
A:Experimental source: strain ALK0243
A:Note: part of the sequence, including the amino end of the mature protein, was conf
C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic
C:Genetics:
A:Gene: phyA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-467/Product: 3-phytase A #status experimental <MAT>
F:27,59,105,120,207,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 72.7%; Score 1688; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 5.2e-132;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY	1	NSHSCDVTVD-GYOC-PEISHLWGOYSPFFSLADESAISDPVPGKCRVTFVQVLSRHGARY 58	
DB	27	NQSTCTDVGQYCFSETSHLWGOYAPFFSLANESAISDPVPGKCRVTFVQVLSRHGARY 86	
QY	59	PTSSSKKYSALIERIQKN-ATFKGKYAFKTYNTYLGADLTFFGENOMVNSGKIFYR 117	
DB	87	PTESKGYKYSALITEEQNVTFDQKGYAFKTYNTYSLGADLTFFGEELVNSGKIFYR 146	
QY	118	YKALNVPVFRASGSDRVIASAKFIEGFSOKLADP---AHQASPVINVIIPGSGY 174	
DB	147	YELRNIIIPFIRSSGSRVIAAGEKIEGFSQSTKLKDPRAQPGSSPKIDVIVSEASS 206	
QY	175	NNTLHGLCTAFEDSTLGDDAEANFTAFAPPIRLE-ALPGVNLTDVNLMDMCF 233	
DB	207	NNTLDPGCTVFEDELADTVANFTAFAPSIORLENDLSGVTLTDTFTYLMDCSF 266	
QY	234	DTVATSDATQSLPCDFTADEW-QYDYLQSL-KYYGYGAGNPLGPAGVGF-NELTAR 290	

Db 267 DTISTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 326
 Qy 291 LTHSPVDQHTSTNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 350
 Db 327 LTHSPVHDDTSSNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 386
 Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLRVNDRVVPLHGGCVD 409
 Db 387 ENITQTDGFSAWTVPFAARLYVEMMOCCA-----EQEPLRVLRVNDRVVPLHGGCPID 439
 Qy 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467
 RESULT 2
 JN0656
 3-Phytase (EC 3.1.3.8) A precursor - Aspergillus niger
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
 C:Species: Aspergillus niger
 C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
 C:Accession: JN0656; S28456
 R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.A.M.J.J.
 Gene 127, 87-94, 1993
 A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)
 A:Reference number: JN0656; MUID:93252284
 A:Accession: JN0656
 A:Molecule type: DNA
 A:Residues: 1-467 <V>AN>
 A:Cross-references: GB:Z16414; NID:g2392; PIDN:CAA78904.1; PID:g2393
 A:Experimental source: strain NRRL3135
 A:Note: parts of the sequence, including the amino end of the mature protein, were confirmed by sequencing
 C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and inorganic orthophosphate from phytate
 C:Genetics:
 A:Gene: phyA
 A:Introns: 15/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phospho
 F:1-23/Domain: signal sequence #status predicted <Sig>
 F:24-467/Product: 3-phytase A #status experimental <MAT>
 F:27, 59, 105, 120, 207, 230, 339, 352, 376, 388/Binding site: carbohydrate (Asn) (covalent) #sta
 F:81, 361/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted
 Query Match 72.6%; Score 1684; DB 1; Length 467;
 Best Local Similarity 73.9%; Pred. No. 1.1e-131;
 Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;
 Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESATSIPDPKGCRTVFFQVLSRHGARY 58
 Db 27 NQSSCDVTVDGQYQCFSETSHLWGOYAPFFSLANESVISPVPAGCRTVFAQVLSRHGARY 86
 Qy 59 PTSSKSKYSALIERIQKNA-TFKGYAFLKTYNTLGGADLTTPFGENQMVNSGKIFYR 117
 Db 87 PTDSKGGKYSALIEIOQNATTPDGKYAFLKTYNYSLGADLTTPFGQELVNSGKIFYR 146
 Qy 118 YKALARNIVPVRASGSDRVTSAREKIEGFSQSAKLADP---AHOASPVINVIPEGSY 174
 Db 147 YESLTRNIVPPIRSNGSRVTSAREKIEGFSQSTKLKDPRAQPGQSSPKIDVISEASS 206
 Qy 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLE-ALPGVNLTDDEDVNNLMDMCPF 233
 Db 207 NNTLDPGTCTVFESELDATVEANFTATFVPSIRQLENLDSGVTLTDTTEVTYLMDCSF 266
 Qy 234 DTVARTSDATQLSPCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
 Db 267 DTISTSTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 326
 Qy 291 LTHSPVDQHTSTNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 350
 Db 327 LTHSPVHDDTSSNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 386

Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLRVNDRVVPLHGGCVD 409
 Db 387 ENITQTDGFSAWTVPFAARLYVEMMOCCA-----EQEPLRVLRVNDRVVPLHGGCPVD 439
 Qy 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467
 RESULT 3
 JN0482
 3-Phytase (EC 3.1.3.8) A - Aspergillus ficum
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
 C:Species: Aspergillus ficum
 C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
 C:Accession: JN0482; PN0023
 R:Ullah, A.H.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 192, 747-753, 1993
 A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemi
 A:Reference number: JN0482; MUID:93249451
 A:Accession: JN0482
 A:Molecule type: protein
 A:Residues: 1-441 <ULL>
 A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
 R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 178, 45-53, 1991
 A:Title: Cyclohexanedione modification of arginine at the active site of Aspergillus
 A:Reference number: PN0023; MUID:91298982
 A:Accession: PN0023
 A:Molecule type: protein
 A:Residues: 48-70 <UL2>
 C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
 F:4, 36, 82, 97, 184, 207, 316, 329, 353, 365/Binding site: carbohydrate (Asn) (covalent) #sta
 F:58, 338/Active site: Arg, His #status predicted
 F:59/Active site: His (phosphohistidine intermediate) #status predicted
 Query Match 71.7%; Score 1665; DB 1; Length 441;
 Best Local Similarity 73.7%; Pred. No. 3.8e-130;
 Matches 328; Conservative 40; Mismatches 59; Indels 18; Gaps 10;
 Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESATSIPDPKGCRTVFFQVLSRHGARY 58
 Db 4 NQSSCDVTVDGQYQCFSETSHLWGOYAPFFSLANESVISPVPAGCRTVFAQVLSRHGARY 63
 Qy 59 PTSSKSKYSALIERIQKNA-TFKGYAFLKTYNTLGGADLTTPFGENQMVNSGKIFYR 117
 Db 64 PTDSKGGKYSALIEIOQNATTPDGKYAFLKTYNYSLGADLTTPFGQELVNSGKIFYR 123
 Qy 118 YKALARNIVPVRASGSDRVTSAREKIEGFSQSAKLADP---AHOASPVINVIPEGSY 174
 Db 124 YESLTRNIVPPIRSNGSRVTSAREKIEGFSQSTKLKDPRAQPGQSSPKIDVISEASS 183
 Qy 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLE-ALPGVNLTDDEDVNNLMDMCPF 233
 Db 184 NNTLDPGTCTVFESELDATVEANFTATFVPSIRQLENLDSGVTLTDTTEVTYLMDCSF 243
 Qy 234 DTVARTSDATQLSPCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
 Db 244 DTISTSTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 303
 Qy 291 LTHSPVDQHTSTNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 350
 Db 304 LTHSPVHDDTSSNHTLSDNATPPLNATLXADSHDNTWYSIFALGLYNGTKPLSTTSV 363
 Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLRVNDRVVPLHGGCVD 409
 Db 364 ENITQTDGFSAWTVPFAARLYVEMMOCCA-----EQEPLRVLRVNDRVVPLHGGCPVD 416
 Qy 410 KLGRCKLDDFVEGLSFARSGGNWAE 434

Db 417 ALGCTRDSFVRLSLARSGDWA 441

RESULT 4

acid phosphatase (EC 3.1.3.2) precursor - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein YAR071W
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 05-May-1995 #sequence_revision 01-Sep-1995 #text_change 20-Apr-2000
 C:Accession: S53476; JCI1018
 R:Bussey, H.; Keng, T.; Storms, R.K.; Vo, D.; Zhong, W.; Fortin, N.; Barton, A.B.; Kabac
 submitted to the EMBL Data Library, February 1994
 A:Description: Sequencing of chromosome I of *Saccharomyces cerevisiae*: analysis of the
 A:Reference number: S53476
 A:Accession: S53476
 A:Molecule type: DNA
 A:Residues: 1-467 <BUS>
 A:Cross-references: EMBL:L28920; NID:g1616966; PIDN:AAC09508.1; PID:g456155; MIPS:YAR071
 R:Chen, J.Y.; Gong, Y.; Ao, S.Z.
 Acta Biochim. Biophys. Sin. 21, 437-444, 1989
 A:Title: The primary structure of acid phosphatase gene PHO11 in *S. cerevisiae* and compa
 A:Reference number: JCI1018
 A:Accession: JCI1018
 A:Molecule type: DNA
 A:Residues: 1-16, 'L', 18-149, 'H', 151-353, 'O', 355-422, 'G', 424-467 <CHE>
 A:Note: this paper is in Chinese, with an English abstract
 C:Genetics:
 A:Gene: SGD:PHO11
 A:Cross-references: SGD:S0000094; MIPS:YAR071W
 A:Map position: 1R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola
 F:1-17/Domain: signal sequence #status predicted <Sig>
 F:18-467/Product: acid phosphatase #status predicted <MAT>
 F:74/Active site: Arg #status predicted
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,162,192,250,315,356,439,445,461/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 15.3%; Score 356; DB 2; Length 467;
 Best Local Similarity 26.9%; Pred. No. 1.3e-21;
 Matches 119; Conservative 63; Mismatches 190; Indels 70; Gaps 17;
 QY 15 EISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALLIERI 74
 Db 36 EIFFPLGGSGPYSPFGDYGISRLDPESCEMKQVMQVGRHGERPTVSKAKSIMTTWYKL 95
 QY 75 QK-NATFGKYAFKL-----TNYNVL-----GADDLTPF-GENQMVNSGIKFYRR 117
 Db 96 SNVTGQFSGLSFLNDYEFFIRDTKNLEMETTLANSVNVNLYPTGEMNAKRHRDFLAQ 155
 QY 118 YKALARNIVPF-VRASGSDRVIAAEKFIQFSAKLADPAHOASPVINVIIPESGYN 176
 Db 156 YGYMVENQTSFAVFTSNRCHDTAQVFIDG-----LQDKFNISLOTISE--AESAGANT 208
 QY 177 TLDHGLCTAFEDSTLGDAAENFTAVFAPPIRRL-EALPGVNLTDDEVNLMDCPFDT 235
 Db 209 LSAHSCPAWDD-VNDLILKKTLYTSLGIAKRLNKENKGLNLTSSDANTFFAWCAYEI 267
 QY 236 VARTSDATQSPFCDLFTADEW-QYDYLQSLK-YGYGAGNPLGPAQGVG-FNELIARLT 292
 Db 268 NAR-----GYSDICNIFTKDELVRFSYGQDLEYQTGPGYDVVRSGANLNFASVKLLK 322
 QY 293 HSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFALGYNKTKPLSTTSVES 352
 Db 323 ESEVQDQ-----KWLSTHTDILNLTITIGIIDDKNLTAHVFPF 364
 QY 353 IETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRVLVNDRVVPLHGGCGVDKLG 412
 Db 365 ME-NTFHRSWYVPOGARVYTERFQC-----SNDTYVRVINDAVVPIETCTSGPGF 414
 QY 413 RCKLDDF-----VEGLSFAR 427

RESULT 6
 PABYCC

Db 415 SCEINDFYDYAEKRVAGTDFLK 436

RESULT 5

acid phosphatase (EC 3.1.3.2) PHO12 - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein YHR215W
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999
 C:Accession: S48996; S59659
 R:Macri, C.
 submitted to the EMBL Data Library, February 1994
 A:Description: The sequence of *S. cerevisiae* cosmid 9177.
 A:Reference number: S46671
 A:Accession: S48996
 A:Molecule type: DNA
 A:Residues: 1-467 <MAC>
 A:Cross-references: EMBL:U00029; NID:g551322; PIDN:AAB69729.1; PID:g458917; MIPS:YHR2
 R:Xu, L.
 submitted to the EMBL Data Library, January 1995
 A:Reference number: S59658
 A:Accession: S59659
 A:Molecule type: DNA
 A:Residues: 1-16, 'L', 18-81, 'AR', 84-149, 'H', 151-467 <XUL>
 A:Cross-references: EMBL:U19789; NID:g847754; PIDN:NAAT3479.1; PID:g847755
 C:Genetics:
 A:Gene: SGD:PHO12
 A:Cross-references: SGD:S00001258; MIPS:YHR215W
 A:Map position: 8R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 15.3%; Score 356; DB 2; Length 467;
 Best Local Similarity 26.9%; Pred. No. 1.3e-21;
 Matches 119; Conservative 63; Mismatches 190; Indels 70; Gaps 17;
 QY 15 EISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALLIERI 74
 Db 36 EIFFPLGGSGPYSPFGDYGISRLDPESCEMKQVMQVGRHGERPTVSKAKSIMTTWYKL 95
 QY 75 QK-NATFGKYAFKL-----TNYNVL-----GADDLTPF-GENQMVNSGIKFYRR 117
 Db 96 SNVTGQFSGLSFLNDYEFFIRDTKNLEMETTLANSVNVNLYPTGEMNAKRHRDFLAQ 155
 QY 118 YKALARNIVPF-VRASGSDRVIAAEKFIQFSAKLADPAHOASPVINVIIPESGYN 176
 Db 156 YGYMVENQTSFAVFTSNRCHDTAQVFIDG-----LQDKFNISLOTISE--AESAGANT 208
 QY 177 TLDHGLCTAFEDSTLGDAAENFTAVFAPPIRRL-EALPGVNLTDDEVNLMDCPFDT 235
 Db 209 LSAHSCPAWDD-VNDLILKKTLYTSLGIAKRLNKENKGLNLTSSDANTFFAWCAYEI 267
 QY 236 VARTSDATQSPFCDLFTADEW-QYDYLQSLK-YGYGAGNPLGPAQGVG-FNELIARLT 292
 Db 268 NAR-----GYSDICNIFTKDELVRFSYGQDLEYQTGPGYDVVRSGANLNFASVKLLK 322
 QY 293 HSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFALGYNKTKPLSTTSVES 352
 Db 323 ESEVQDQ-----KWLSTHTDILNLTITIGIIDDKNLTAHVFPF 364
 QY 353 IETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRVLVNDRVVPLHGGCGVDKLG 412
 Db 365 ME-NTFHRSWYVPOGARVYTERFQC-----SNDTYVRVINDAVVPIETCTSGPGF 414
 QY 413 RCKLDDF-----VEGLSFAR 427
 Db 415 SCEINDFYDYAEKRVAGTDFLK 436

271 --GYSDVCIDTFEDELVRYSYGQDLVSFYQDGPYDMIRSVGANLNFATLKLKXSEIQD 320
299 HTSTNHTLSDNSPATFPLNATLYADSHDNTWYSIFFALGLYNGTKPLSTTSVESIEDGY 358
329 -----LKWLSFTHDTHDILNLTAGIIDDKNLTAEYVPFM-GNTF 369
359 AASWTVPFAARAAYVENMOCEAGGGGEGEKEPLRVVLVNDVRVPLHGGGVDKLGRCKLDD 418
370 HKSWYVPGARVYTEKFQC-----SNDTVRVYVINDAVVPIETCSTGPGFSCEIND 420
419 F-----VEGLSFAR 427
421 FDYAEKRVAGTDFLK 436
RESULT 7
PABYC
acid phosphatase (EC 3.1.3.2) repressible, precursor - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: acid phosphatase PHO5; protein YBR0814; protein YBR093c
C:Species: *Saccharomyces cerevisiae*
C:Date: 19-Feb-1984 #sequence revision 30-Sep-1991 #text change 12-Nov-1999
C:Accession: S05795; A38792; S48260; S45961; A00777; A38793; S41855; B25241; A25367;
R:Baizha, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
Nucleic Acids Res. 12, 7721-7739, 1984
A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in *y*
A:Reference number: S05794; MUID:85037940
A:Accession: S05795
A:Molecule type: DNA
A:Residues: 1-467 <BAJ>
A:Cross-references: EMBL:X01079; NID:g4162; PIDN:CAA25555.1; PID:g758282
A:Note: the authors translated the codon TAC for residue 272 as Thr
A:Accession: A38792
A:Molecule type: protein
A:Residues: 18-45 <BAJ2>
R:Manhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.
Yeast 10, 1363-1381, 1994
A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A:Reference number: S48255; MUID:95208357
A:Accession: S48260
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-467 <MAN>
A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55598.1; PID:g476051
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R:Feldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45927
A:Accession: S45961
A:Molecule type: DNA
A:Residues: 1-467 <PE2>
A:Cross-references: EMBL:X235962; NID:g536364; PIDN:CAA85046.1; PID:g536365; GSPDB:GNO
R:Arima, K.; Oshima, T.; Kubota, I.; Nakamura, N.; Mizunaga, T.; Toh-e, A.
Nucleic Acids Res. 11, 1657-1672, 1983
A:Title: The nucleotide sequence of the yeast PHO5 gene: a putative precursor of repr
A:Reference number: A00777; MUID:83168913
A:Accession: A00777
A:Molecule type: DNA
A:Residues: 1-35, 'Y', 37-129, 'G', 131-293, 'O', 295-445, 'V', 447-461, 'DT', 464-465, 'K', 467
A:Cross-references: EMBL:V01320; NID:g4158; PIDN:CAA24630.1; PID:g4159
A:Accession: A38793
A:Molecule type: protein
A:Residues: 18-26, 'X', 28 <ARI2>
R:Meyhack, B.; Baizha, W.; Rudolph, H.; Hinnen, A.
EMBO J. 1, 675-680, 1982
A:Title: Two yeast acid phosphatase structural genes are the result of a tandem dupli
A:Reference number: S41855; MUID:84236032
A:Accession: S41855
A:Molecule type: DNA
A:Residues: 1-50, 'T', 32-51, 'S', 53-75 <MEY>
A:Cross-references: EMBL:M24178; NID:g172156; PIDN:AAA34868.1; PID:g172157
R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill
Mol. Cell. Biol. 6, 1855-1865, 1986

[illegible]

A:Reference number: A93074; MUID:87064474

A:Accession: B25241

A:Molecule type: DNA

A:Residues: 1-44 <TAI>

R: Bergman, L.W.

Mol. Cell. Biol. 6, 2298-2304, 1986

A:Title: A DNA fragment containing the upstream activator sequence determines nucleosome

A:Reference number: A25367; MUID:87064526

A:Accession: A25367

A:Molecule type: DNA

A:Residues: 1-2, 'Y', 'A', '4-43, 'T', '45-51 <BER>

R: Silver, S.; Monod, M.; Hinnen, A.; Hagnenauer-Tsapis, R.

Mol. Cell. Biol. 7, 3306-3314, 1987

A:Title: The yeast acid phosphatase can enter the secretory pathway without its N-termin

A:Reference number: A27774; MUID:88038886

A:Accession: A27774

A:Molecule type: DNA

A:Residues: 1-51, 'S', '53-60 <STL>

A:Cross-references: GB:M17306

C: Genetics:

A:Gene: SGD:PHO5; MIPS:YBR093c

A:Cross-references: SGD:S0000297; MIPS:YBR093c

A:Map position: 2R

A:Note: YBR093c

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphatidine; phosphoprotein; phosphoric monoester hydrola

F.1-17/Domain: signal sequence #status predicted <SIG>

F.18-467/Product: acid phosphatase, repressible #status experimental <MAT>

F.75/Active site: His (phosphatidine intermediate) #status predicted

F.97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (coval

F.337/Active site: His #status predicted

Query Match 15.0%; Score 347; DB 1; Length 467;
Best Local Similarity 26.5%; Pred. No. 7.5e-21;
Matches 117; Conservative 57; Mismatches 187; Indels 80; Gaps 18;

QY 21 GQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTSSKSKYSALIERIQKNAT 79

Db 42 GGAGPYVSPGDYIGSRDLPECEMKQLQWGRGERIPTVSLAKTIKSTWYKLSNTRQ 101

QY 80 FKGVAFLEK-TNYTLGND-----LTPF-GENQMVNSGKIFRYRYKALAR 123

Db 102 FNGSLFNDYEFFIRDDDDLEMETTFANSDDVNLPTGEMNAKRHARDPLAQGYAVE 161

QY 124 NIVPF-VRASGDRVIAAEKFIQFQSLADPAHQASPVINII-----PEGGYNN 177

Db 162 NOTSFVFTSNKRCHDTAQYFIDG-----LGDO-----FNITQTVSEASAGANTL 209

QY 178 LDHGLCTAFEDSTLGDAAEFNFAVAPPFIRARL-EALPGVNLTDQVNLMDMCPFDTV 236

Db 210 SACNSCPAW-DYDANDDIVNEYDTTLDIAKRLNKENKGLNTSTDASTLFSWCAFEV- 267

QY 237 ARTSDATQSPFCDLETADSW-QYDYLOSL-KYGYGAGNPLGPAQGVG-FNELIARLTH 293

Db 268 -----NAKGYSDVCDIFTKDELVHYQYQDLHTYHCGPGYDIITKSGNLFNASVKLLQ 323

QY 294 SPVQDHTSNHTLDSNPATFPLNATLYADFSDNTMVSIFFAFGNGTKPLSTTSVESTI 353

Db 324 SETQDQ-----KWLSTHTDILNLTLAGIDDKNNLTAEVVPFM 365

QY 354 ETDGYAASWTPFAARAYVEMMOCEAGGGGEGEKEPVRVLVNRVPLHGGGVKGLGR 413

Db 366 -GNTFHRWYVPGGARVYTEKFC-----SNDTYVRYVINDAVVPIETCSTGPGFS 415

QY 414 CKLDDE-----VEGLSFAR 427

Db 416 CEINDFYDAEKRVAGTDFLK 436

RESULT 8

S52495

acid phosphatase homolog YDL024c - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein D2815

C:Species: Saccharomyces cerevisiae

C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000

C:Accession: S52495; S67556

R:Andre, B.; Viissers, S.; Urrestazu, L.

submitted to the EMBL Data Library, February 1995

A:Description: The sequence of a 42 kb segment located on the left arm of chromosome

A:Reference number: S52492

A:Accession: S52495

A:Molecule type: DNA

A:Residues: 1-468 <AND>

A:Cross-references: EMBL:Z48432; NID:9683669; PIDN:CAA88335.1; PID:9683673

R:Urrestazu, L.A.; Andre, B.; Viissers, S.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67533

A:Accession: S67556

A:Molecule type: DNA

A:Residues: 1-468 <URR>

A:Cross-references: EMBL:Z74072; NID:91430996; PIDN:CAA98583.1; PID:g1430997; MIPS:YD

A:Experimental source: strain S288C

C: Genetics:

A:Map position: 4L

C:Superfamily: yeast acid phosphatase

Query Match 14.5%; Score 337; DB 2; Length 468;
Best Local Similarity 26.0%; Pred. No. 5e-20;
Matches 119; Conservative 60; Mismatches 161; Indels 118; Gaps 19;

QY 21 GQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTSSKSKYSALIERIQKNAT 80

Db 43 GGSADYFSPFANYGPTDIPGCRLLTQVMIGRHGRIYTRSEAKD---IFEVWKISNY 99

QY 81 KGKY----AFLEKT-YNYTL-----GADDLTPF-GENQMVNSGKIFRYRYKA 120

Db 100 TKGYBSLSFLNNGYEFPIDESLLEMETTLQNSIDVNLPTGEMNAKRHAREFLAKYKG 159

QY 121 LARNIVPF-VRASGDRVIAAEKFIQFQSLADPAHQASPVINIIPEGSGVNNITLD 179

Db 160 LMENCTNPFITTSKRIYDTAQVFAEAL-----DGFNLSIQ 197

QY 180 HGLCTAFEDSTLGD-----DAAENATVAFAPPFIRARL-EALPGVNL 220

Db 198 ----TLSENSSGANTLAAKSSCPNWSNANNILMSYSDYLENISDRNDENKGLNLS 253

QY 221 DEDVNNMDMCPFTVARTSDATQSPFCDLETADSW-QYDYLOSL-KYGYGAGNPLGP 278

Db 254 RKDAALFSCAFEL-----NAKGYSDVCDIFSAELIHYSETDLTSFYQNGPGYKLIK 308

QY 279 AQGVG-FN---ELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFSDNTMVSIF 334

Db 309 SIGANLENATVKLIQSAH-----LDQKWLSTHTDILNLT 347

QY 335 ALGYNGTKPLSTTSVESTIDGYAASWTPFAARAYVEMMOCEAGGGGEGEKEPVRV 394

Db 348 TAGLIDDTNLTNNHV-PFRDHSYHRSWYIPOGARVYTEKFC-----SNDSYVRY 397

QY 395 LVNDRVYVPLHGGGVKGLGRCKLDDEFE-----GLSF 425

Db 398 VNDVAVPIESCSSGGFSCCEGTFEYAKRLRGVSF 435

RESULT 9

JC4285

acid phosphatase (EC 3.1.3.2) precursor - yeast (Pichia pastoris)

N:Alternate names: orthophosphoric monoester phosphohydrolase (acid optimum); Pho

C:Species: Pichia pastoris

C>Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999

C:Accession: JC4285

R:Payne, W.E.; Gannon, P.M.; Kaiser, C.A.

Gene 163, 19-26, 1995

A:Title: An inducible acid phosphatase from the yeast Pichia pastoris: Characterizati

Sat Oct 27 15:25:43 2001

C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-479/Product: 3-phytase #status predicted <MAT>
 F:81.337/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted
 F:106,191,227,250,315,340,425,442,458/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 13.28; Score 307.5; DB 1; Length 479;

Best Local Similarity 27.48; Pred. No. 1.5e-17;
 Matches 124; Conservative 55; Mismatches 175; Indels 99; Gaps 23;

QY 9 DGYQCPETSHLWQ--YSPFFSLADESAISDPKGRVTFVQVLSRHGARYPTSSSKK 66
 DB 41 DGYSI--LKHGNGPYSERSV---GIARDPPTSCEVDQVIMVKRGERYPSPSACKG 94

QY 67 YSALIERIOK-NAT-FKGYAFKTYNYTL-----GADDLT-PF-GENQMVNSIKFYR 116

DB 95 IEALAKVYSINTTEYKGLAFNDWTYVPNECYNAETTSYAGLLDAYNHGNDYKA 154

QY 117 RYKAL--ARNIVPFVRASGSDRVASAEKFIQFQSAKLADPAHQASPVINVIPEGSGY 174

DB 155 RYGLHWGETVVPFF--SSGYGRVETARKEGSGF-----FCY 190

QY 175 NNTLDHGLCTAFEDSTLGDDAE-----NFTAVFAPPIR---ARLEAL-PGV 217

DB 191 NYSNAALNIISEVMGADSLTPTCDTNDQTCNDLTQYL-POFKVAARLNSONPOM 249

QY 218 NLTDDEVNLMDCPDTVARTSDATQLSFCDLFTADEW-QYDYLQSLKYGYGAG--- 273

DB 250 NLTAADSYNLMWMAFELNAR-----PFSNWINAFQDEWVSFGYVEDLNY--YCAGPGD 303

QY 274 NPLGPAQGVGFNELLIALRTHSPVDHTSTNHTLDSNPATPLNATLYADESHDNTWVSIF 333

DB 304 KNMAAVGANVYANASLTLLNQPKB-----AGSLFFNFHAHDTNTPIL 345

QY 334 FALGLY--NGTKPLSTSVESIEFTDGYAASWTVPFAARAYVEMMQCEAGGGEKEPL 391

DB 346 AALGVLPINEDLPDRVAF-----GNPYSIGNIVPMGHLITIERLSQATALSDEG---TY 398

QY 392 VRVLVNDVRVPLHGGVDBKLGRCCKLDDFVEGLS 424

DB 399 VRVLVNEAVLPFNDCTSGPGYSCPLANTYTSILN 431

RESULT 11

JN0715

3-phytase (EC 3.1.3.8) B precursor - Aspergillus ficuum

N:Alternate names: pH 2.5-optim acid phosphatase

C:Species: Aspergillus ficuum

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

C:Accession: JN0715; PNO594; PNO460

R:Erlich, K.C.; Montalbano, B.G.; Mullaney, E.J.; Dischinger Jr., H.C.; Ullah, A.H.J

Biochem. Biophys. Res. Commun. 195, 53-57, 1993

A:Title: Identification and cloning of a second phytase gene (phyB) from Aspergillus

A:Reference number: JN0715; MUID:93371452

A:Accession: JN0715

A:Molecule type: DNA

A:Residues: 1-479 <EHR>

A:Cross-references: GB:L20567

A:Accession: PNO594

A:Molecule type: protein

A:Residues: 20-101;133-146;376-399 <EH2>

R:Ullah, A.H.J.; Dischinger Jr., H.C.

Biochem. Biophys. Res. Commun. 192, 754-759, 1993

A:Title: Identification of active-site residues in Aspergillus ficuum extracellular p

A:Reference number: PNO460; MUID:93249452

A:Accession: PNO460

A:Molecule type: protein

A:Residues: 65-66,68-93 <ULL>

C:Comment: This enzyme, previously characterized as an acid phosphatase (EC 3.1.3.2),

C:Genetics:

A:Reference number: JC4285; MUID:96001238

A:Accession: JC4285

A:Molecule type: DNA

A:Residues: 1-468 <PAY>

A:Cross-references: GB:U28658; NID:9881955; PIDN:AAA85503.1; PID:9881956

A:Experimental source: GS115

C:Genetics:

A:Gene: phol

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-468/Product: acid phosphatase #status predicted <MAT>

F:84/Active site: His (phosphohistidine intermediate) #status predicted

F:163,196,256,321,360,435/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:345/Active site: His #status predicted

Query Match 14.48; Score 334.5; DB 2; Length 468;

Best Local Similarity 27.08; Pred. No. 8.1e-20;

Matches 124; Conservative 65; Mismatches 187; Indels 83; Gaps 22;

QY 7 TVDGYQCPETSHLWQYSPF-----FSLADESAISDPKGRVTFVQVLSRHGARYPTS 61

DB 40 TDDQYNI--LRHL-GGLGPGYCYNGWGAEESEI-----ESCTIDQAHLLMRHGERYPST 91

QY 62 SKSKYSALIERIOKNATFK---GKYAFKTYNYTLG-----ADLTP---FGENQMVNS 110

DB 92 NVGKQLEALYQKL-LDADVEVPTGPIFFQDYDFVSDAAWYEQETTKGFYSGLTAPDF 150

QY 111 GIKFYRYKALARNIVP-----FVRASGSDRVASAEKFIQFQSAKLADPAHQASPVINV 166

DB 151 GTTLREYDHLINTSEEGKLSWAGSQERVDYTAQAGFMKKNYTD-----WVEV 203

QY 167 IIPF--GSGYNTLDHGLCTAFEDSTLGDDAEANFTAVFAPPIRARLEAL-PGVNLDE 222

DB 204 VALEEKSGSLNLTARISCPNPNVSHYKD---GDFPNDAIERADRLNTLSPGFENITAD 260

QY 223 DVNLMDCPDTVARTSDATQLSFCDLFTADEWQYD-YLQSLK-YCYGAGNPLGPAQ 280

DB 261 DIFTIALYCGFELNVRGE-----SSPCDLSREALLYTAYLRDLGWYNNVNGNPLGKTI 315

QY 281 GVGFNELIARLTHSPVDHTSTNHTLDSNPATPLNATLYADESHDNTWVSIFFAFALGLYN 340

DB 316 GYVYANATROLLEN-----TEADPRDP-----LYFSFSDTDLLOVFTSLGLEN 360

QY 341 GTK-PLS---TTSVESIETDGYAASWTVPFAARAYVEMMQCEAGGGEKEPLVRVL 395

DB 361 VTDLPDQIQFQTSFKSTE-----IVPMGARLLTERLLCTV-----EGEEKYVVRTI 407

QY 396 VNRVPLHGGVDBKLGRCCKLDDFVEGLSFARSGGNWAE 434

DB 408 LNDVAVPLSDCSGPGFSCPLNDYVSRLEALNEDSDFAE 446

RESULT 10

JN0890

acid phosphatase (EC 3.1.3.2) precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0890

R:Pidington, C.S.; Houston, C.S.; Palohelmo, M.; Cantrell, M.; Miettinen-Oinonen, A.; N

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-optim

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0890

A:Molecule type: DNA

A:Residues: 1-479 <PID>

A:Cross-references: GB:L02420; NID:g166481; PIDN:AAA16897.1; PID:g166482

A:Experimental source: strain ALK0243

C:Comment: The highly similar enzyme from A. ficuum has been shown to have 3-phytase (EC

C:Genetics:

A:Gene: aph

A:Introns: 261/1; 300/2; 335/2

C:Superfamily: yeast acid phosphatase
C:Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:68/Active site: Arg #status predicted
F:69/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 12.8%; Score 298; DB 1; Length 453;
Best Local Similarity 25.6%; Pred. No. 8.2e-17;
Matches 107; Conservative 63; Mismatches 182; Indels 66; Gaps 16;

QY 23 YSPFFSLADESAISPDVPKGRVTFQVQLSRHGARYPTSSKKSKYSALIERIQK----NAT 79
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 43 HKPYE---YGPSIDFPPTCKIKQVHTLQRGSRNPCTGGNAADFAGTANFOORLLNGS 97
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 80 FKGKY-----AFUKTYNTL---GADDLTPFGENQMVNSGIKFYRYRKALARNIVPV 129
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 98 VPIDSVSGNPLSFSTWPTVPVEAANADALSRRGVLEFDMGQRQFYERHFLFNASTYNI 157
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 130 RASGSDRVIASAEKTIEGFQSACLADPAHQAPVINVIPEGSGVNTNLDHGLTAPEDS 189
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 158 YTAQAQRVVDSALWCYGM---FCEDVHNFTNYILVSENATAGSNLSYINACPASDAD 213
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 190 TLGDDEANFTAVAPPTRIRLEA-LPGVNLTDEDVNVLMDCMPDITVARTSDATQLSPF 248
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 214 DFTTPALEAWRNVMPPRIQRNLPNPFSYNLTNDILNLGYCSYEIALQ-----DYSEF 268
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 249 CDLF-TADEWOVDYLQSLKYYCYGAGNPL--CPAGCVGNELIARLTHSPVODHFTSTNHT 305
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 269 CKLENSVDFLNFVEGDL-SYGMGNSVKWSIGFPGAAYANSLANSLRS-VENNTHQ---- 322
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 306 LDSNPATPLNATIYDFSHDNTMVSIFFALGLYNGT---KPLSTTSVESIETDGAAASW 362
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 323 -----QVFFAPTHANIIPVELTGFTDNTPENPLPS--YQHISMKASE 368
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 363 TVPTAARAAYVEMMOCEAGGGGGEKEPLVRVLVNDVRVPLHGCVGDKL----GRCKL 416
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 369 FVPAGNLI TELFOE-----DSKYVYRHLYNEEVFPLSDCGFGFNTSDGMCEL 418
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13
c141119
acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
;;Species: Schizosaccharomyces pombe
;;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
;;Accession: S14119; T40455
;;Yang, J.; Schweingruber, M.E.
;;Eur. J. Genet. 18, 269-272, 1990
;;Title: The structural gene coding for thiamin-repressible acid phosphatase in
;;Reference number: S14119; MUID:91064763
;;Accession: S14119
;;Status: preliminary
;;Molecule type: DNA
;;Residues: 1-463 <YAN>
;;Cross-references: GB:X56939; NID:g5006; PIDN:CAA40258.1; PID:g5007
;;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.
Experimental submitted to the EMBL Data Library, November 1998
Reference number: Z21931
Accession: T40455
Status: preliminary; translated from GB/EMBL/DDBJ
Molecule type: DNA
Residues: 1-463 <LYN>
Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN00067; SPDB:SPBC428.0
Experimental source: strain 972h; cosmid c428
Genetics:
Gene: SPBC428.03c
Map position: 2
Superfamily: yeast acid phosphatase
Keywords: phosphoric monoester hydrolase

Query Match 12.5%; Score 290; DB 2; Length 463;
Best Local Similarity 24.3%; Pred. No. 3.9e-16;
Matches 109; Conservative 70; Mismatches 183; Indels 86; Gaps 18;

Db 212 VFQDNFHNKATDAAHAYVRNIFIEPIVNRLLAKYFDSSYKLTINDVRSIFYCEIEIAK 271
 QY 239 TSDATQLSDFCDLFTADEW-QYDYLOSLKYYGYGAGNPLGPAQGVGFENLIARLTHSPVQ 297
 Db 272 DH-----SDFCSIFTPSEFLNFEYDSDL-QAYGG-----PVS 304
 QY 298 DHTST-----NHTLDS--NPATFPLNATLYADFSDHNTMWSIFFALGLY---NCTKPLS 346
 Db 305 EWASTLGGAYINLADSLRNVTNPFDRKVFLLAFTHDSNLIIPVEAALGFFPDITPONPLP 364
 QY 347 TTSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLYNDRVPLHCC 406
 Db 365 TD--KNLYTSQKTSFVFPAGNLITELFFC-----SDSKYYVRHLVNOOVYPLIDC 414
 QY 407 GVDKLGK-----CKLDDFVBLSPFARSGGNWAEFC 436
 Db 415 GYGPSGTSGLCLCQLQAYLNSPIRANSTNSNGISIF 448

RESULT 15

A86233
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A86233
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzia
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: A86233
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-468 <S>TO>
 A:Cross-references: GB:AF005172; NID:g2150177; PIDN:AAB60740.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 6.6%; Score 154; DB 2: Length 468;
 Best Local Similarity 22.9%; Pred. No. 7.4e-05;
 Matches 96; Conservative 52; Mismatches 184; Indels 88; Gaps 18;

QY 38 DVPKGRVTFVQVLSRHGARYPTSSKSKYSALIERI-----Q 75
 Db 50 NVPSECTPIHLNVARHGTSPTRKRLRELESAGRFKELVDAEARKLPDSKIPGWLQ 109
 QY 76 KNATFGKYAFKTYNTLTGADDLTPFGENOMVNSGIKRYRYKALARN-----IVPFVRA 131
 Db 110 WSPWEGK-----VKGELIRQGEDELYQLGIRVRFRFSLFEEDYHPDYITIRA 159
 QY 132 SGSDRVIAAEKFEQFSQAKLADPAHQASPVINVLIEPGSGYNNLTHGLCTAPEDSTL 191
 Db 160 TQTPRASASAVAFGMGLFSEK-----GNL-----GPGNRFAF-----AVTSEN 198
 QY 192 GDAAENFTAVFAPPIRARLEALPGVNLTDVNLMDMC-----PFDVTARTSDATQLSPF 248
 Db 199 ASDTKLRFEECCONYSKYKAKEPAVDKLPVNLKITAASVAKRYDLKFTKODISLWFL 258
 QY 249 CDLFTADEWQYDYLOSLKYYGYGAGNPLGPAQGVGFENLIARLTHSPVQDHTSTNHTLDS 308
 Db 259 CKQVALLWT-DDLEVEFLKGY--GNSLNYKMGV---PULLEDVLHSMEEAIAKREKLP- 311
 QY 309 NPATFPLNATLYADFSDHNTMWSIFFALGLY-NGT-----KPLSTTSVESIETDGYA 359
 Db 312 -PGSYE-KARL--RFAHAETIVFSCLLGLFDGSEFEKIQREKPLELPP-OPPKTRDRF 366

QY 23 YSPFSLADESAISPDVPKGRVTFVQVLSRHGARYPTSS-----KSKYSALIERIQK 76
 Db 43 HEPYF-----NGPTTSPESCAIKQVHLLQHGSRNPTGDDTATDVSAQYIDIFQNKLL 97
 QY 77 NATFGKYA-----FLKTYNTLT--GADLTPFGENOMVNSGIKRYRYKALARNIV 126
 Db 98 NGSIPIVNFSEYENPLFYFKHPTVPVKAENADOLSSGRILELFDLGRQVFEYELFDTDV 157
 QY 127 PVRASGSDRVIAAEKFEQFSQAKLADPAHQASPVINVLIEP--SGYNNLTHGLCT 184
 Db 158 YDINTAAQERVVDSAEWFSYGMFGDDMNKTN-----FVLPEDDSGAGANSLAMYS 211
 QY 185 AFEDSTLGGD-----AEANFTAVFAPPIRARLEAL--PGVNLTDVNLMDMCDFDTVAR 238
 Db 212 VYEDNNIDENTTAAHTSWRNVFLKPIANRLNKYFDGSGYNTLTVSDVRSLYIYICVETALR 271
 QY 239 TSDATQLSDFCDLFTADEW-QYDYLOSLKYYGYGAGNPLGPAQ-----GVGFENLIAR 290
 Db 272 DN-----SDFCSIFTPSEFLNFEYDSDLQAYWG-----GPASEWASTLGGAYVNNLANN 321
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMWSIFFALGLY---NCTKPLST 347
 Db 322 L-----RKGNNASDRK-----VFLAFTHDSQIIPVEAALGFFPDITPEHPLPT 365
 QY 348 TSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLYNDRVPLHCCG 407
 Db 366 D--KNLYTSQKTSFVFPAGNLITELFFC-----SDKNYYVRHLVNOOVYPLIDCG 415
 QY 408 VDKLGR-----CKLDDFVBLSPFARSGGN 431
 Db 416 YGPGSAGSLCGLCALSAYLNSSVRYNSTSN 443

RESULT 14

T39929
 thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39929
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21857

A:Accession: T39929

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-463 <LYN>

A:Cross-references: EMBL:AL023286; PIDN:CAAL18863.1; GSPDB:GN00067; SPDB:SPBC21H7.03c

A:Experimental source: strain 972h; cosmid c21H7

C:Genetics:

A:Gene: SPDB:SPBC21H7.03c

A:Map position: 2

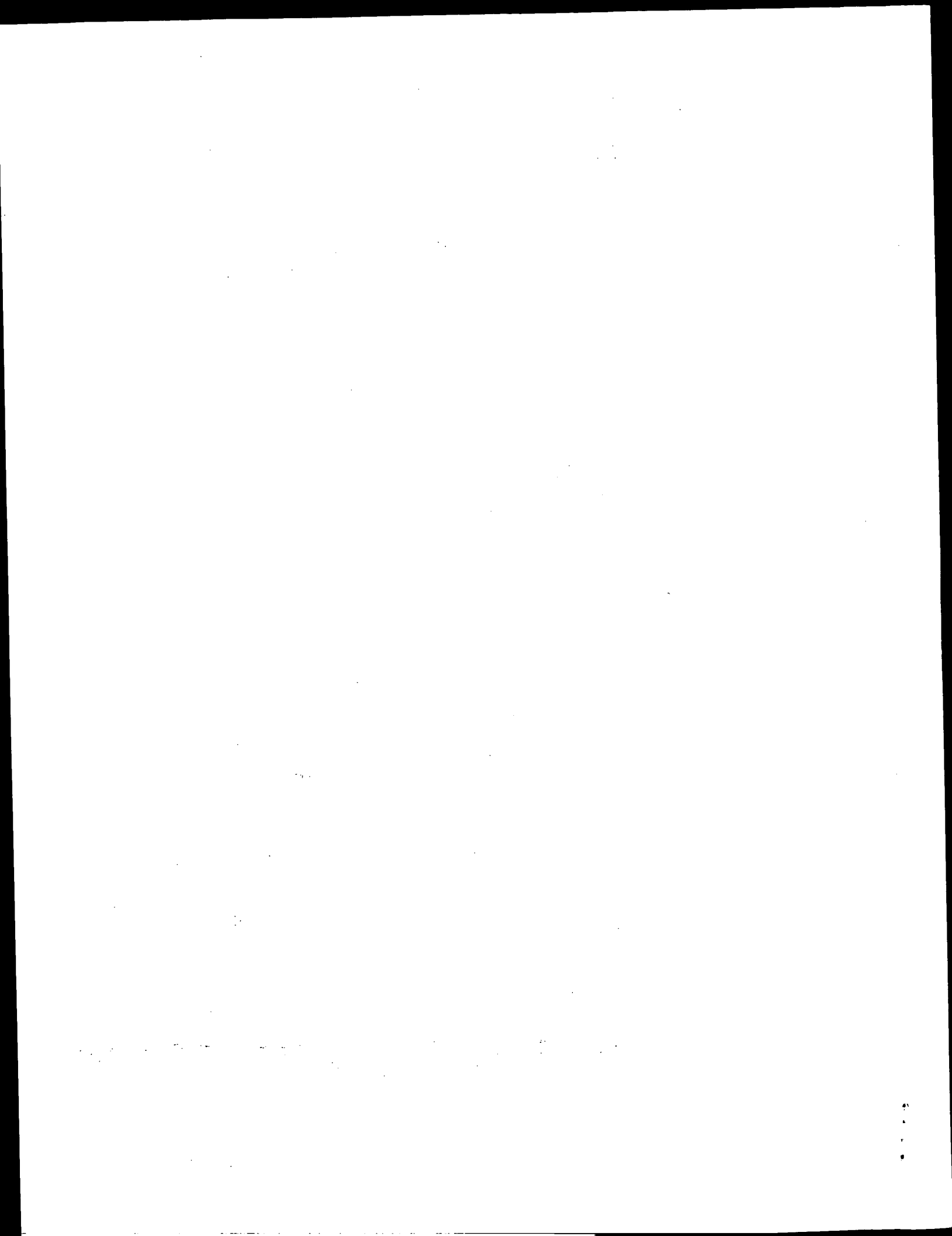
C:Superfamily: yeast acid phosphatase

Query Match 12.3%; Score 285; DB 2: Length 463;
 Best Local Similarity 24.9%; Pred. No. 1e-15;
 Matches 113; Conservative 69; Mismatches 184; Indels 88; Gaps 19;

QY 23 YSPFSLADESAISPDVPKGRVTFVQVLSRHGARYPTS--SKSKYSA-----L 70
 Db 43 HEPYFDGLDSA-----FPETCEIQVHLLQHGSRNPTGDDTATDVYSSQYNNFQEKLL 97
 QY 71 IERIQKATFP-KGKYAFKTYNTLT--GADLTPFGENOMVNSGIKRYRYKALARNIV 126
 Db 98 NGSIPIVNFSEYENPLCYFTQWTPVDAENADOLSSGRILELFDLGRQYKLYKLFDSYV 157
 QY 127 PVRASGSDRVIAAEKFEQFSQAKLADPAHQASPVINVLIEP--SGYNNLTHGLCT 184
 Db 158 YDINTAAQERVVDSAEWFSYGMFGDDMNKTN-----FVLPEDDSGAGANSLAMYS 211
 QY 185 AFEDSTL-----GDDAEANFTAVFAPPIRARLEAL--PGVNLTDVNLMDMCDFDTVAR 238

Qy 360 ASWTVPFAARAYVENMOCEAGGGGEGEKEPLVRVLYNDRVVPLHGCGVDKLGRCCKLDDF 419
Db 367 GSTMAPFGGNNILVLYSCPA-----ESSPKYFVQVLHNEHPVAVPGC--DGKDFCPLDF 419

Search completed: October 26, 2001, 15:10:28
Job time: 6334 sec



Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1688	72.7	467	1	PHYA_ASPPAW	P34753	aspergillus
2	1684	72.6	467	1	PHYA_ASPPNG	P34752	aspergillus
3	1676.5	72.2	463	1	PHYB_EMENI	Q00093	emerigella
4	356	15.3	467	1	PPAB_YEAST	P35842	saccharomyc
5	356	15.3	467	1	PPAC_YEAST	P38693	saccharomyc
6	353	15.2	467	1	PPA3_YEAST	P24031	saccharomyc
7	347	15.0	467	1	PPA5_YEAST	P00635	saccharomyc
8	337	14.5	468	1	PPAD_YEAST	P52290	saccharomyc
9	334.5	14.4	468	1	PPAL_PICPA	P52291	pichia past
10	307.5	13.2	479	1	PHYB_ASPPAW	P34755	aspergillus
11	305.5	13.2	479	1	PHYB_ASPPNG	P34754	aspergillus
12	300.5	12.9	469	1	PPA5_KLULA	P52289	kluyveromyc
13	298	12.8	453	1	PPAL_SCHPO	P08091	schizosacch
14	290	12.5	463	1	PPA2_SCHPO	Q01682	schizosacch
15	125	5.4	755	1	PPAX_CAEEL	Q09549	caenorhabdi
16	113.5	4.9	423	1	PPAL_RAT	P20611	rattus norv
17	111	4.8	693	1	LYSA_YEAST	P49367	saccharomyc
18	108.5	4.7	421	1	PPAL_MOUSE	P24638	mus musculus
19	107.5	4.6	381	1	PPAP_RAT	P20646	rattus norv
20	104.5	4.5	413	1	PPAW_CAEEL	Q09451	caenorhabdi
21	104.5	4.5	423	1	PPAL_HUMAN	P11117	homo sapien
22	102	4.4	1081	1	MSH3_ARATH	Q05607	arabidopsis
23	98.5	4.2	341	1	AAFP_RHILV	Q52812	rhizobium l
24	98.5	4.2	583	1	CYSP_PLAVI	P42666	plasmodium
25	98.5	4.2	1527	1	MRP3_HUMAN	Q15438	homo sapien
26	97.5	4.2	392	1	G3PA_TOBAC	P09043	nicotiana t
27	97	4.2	1458	1	PHLX_RABIT	Q05017	oryctolagus
28	96.5	4.2	844	1	YD25_SCHPO	Q03712	schizosacch
29	96.5	4.2	1350	1	VG72_HSVII	Q00103	ictalurid h
30	96.5	4.2	2290	1	POUG_EMCV	P03304	encephalomy
31	95.5	4.1	394	1	PGK_BACME	P24269	bacillus me
32	94	4.0	776	1	ISOA_PSEAY	P10342	pseudomonas
33	93.5	4.0	776	1	ISOA_PSEAY	P26501	pseudomonas

DISULFID	215	465	BY SIMILARITY
DISULFID	264	282	BY SIMILARITY

ACT_SITE	82	82	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
ACT_SITE	361	361	PROTON DONOR (BY SIMILARITY).
DISULFID	31	40	BY SIMILARITY.
DISULFID	71	414	BY SIMILARITY.
DISULFID	215	465	BY SIMILARITY.
DISULFID	264	282	BY SIMILARITY.
DISULFID	436	444	BY SIMILARITY.

Sat Oct 27 15:25:43 2001

FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 467 AA; 51075 MW; 1186828A5D/EC661 CRC64;

Query Match 72.7%; Score 1688; DB 1; Length 467;
 Best Local Similarity 73.9%; Pred. No. 1.6e-132;
 Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDTVD-GYQC-PEISHLWQYSPFFSLADESAISPDVPGKCRVTFVOVLSRHGARY 58
 DB 27 NOSTCDTVDGQYCFSETSHLWQYAPFSLANESAISSPDVPGKCRVTFVOVLSRHGARY 86
 QY 59 PHSKSKYSALIERIOKN-ATEKGYAFKTYNTLTGADDLTPFGENOMVNSGKIFRR 117
 DB 87 PTESGKKYSALIEELIQNVTFDGYAFKTYNTLTGADDLTPFGEQELVNSGKIFQYR 146
 QY 118 YKALARNIVPFRASGDSRVIASAEKFTGFSQAKLADP---AHOQSPVINVIPEGSY 174
 DB 147 YESLTRNIIPFRSSGSRVIAAGEKFTGFSQAKLADP---AHOQSPVINVIPEGSY 206
 QY 175 NNTLDHGLCTAFEDSLGDAEANTAVFAPPIRLE-ALPGVNLTDDEVDVNLMDMCP 233
 DB 207 NNTLDHGLCTAFEDSLGDAEANTAVFAPPIRLE-ALPGVNLTDDEVDVNLMDMCP 266
 QY 234 DTVARTSDATQSPFCDFLETADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELTAR 290
 DB 267 DTISTSTVDTKLSPFCDFLETHDEWIHYDYLQSLKYYGAGNPLGPAQGVGF-NELTAR 326
 QY 291 LTHSPVODHTSTNHTLDSNPATFLNATYADSHNTMVSIFPFGALGLYNGTKPLSTTSV 350
 DB 327 LTHSPVODHTSTNHTLDSNPATFLNATYADSHNTMVSIFPFGALGLYNGTKPLSTTSV 386
 QY 351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLRVVLVNDVPLHCGGVD 409
 DB 387 ENITQTDGFSAMTVPASRLYVEMMQCEAGGGGEGEKEPLRVVLVNDVPLHCGGVD 439
 QY 410 KLGRCCKLDDFVEGLSFARSGNNAECPA 437
 DB 440 ALGRCTRDSFVRGLSFARSGDNWAECSA 467

RESULT 2

PHYA_ASPNG STANDARD; PRT; 467 AA.

AC P34752;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE A).
 GN PHYA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX STRAIN-NRRL 3135 / VAN TIEGHEM / FICUUM;
 RX MEDLINE=93252284; PubMed=8387447;
 RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
 RA Gouka R.J., Suykerbuijk M.E.G., Luiten R.G.M., van Paridon P.A.,
 RA Seltén G.C.M., Veenstra A.E., van Gorcom R.F.M.,
 RA van den Hondel C.A.M.J.J.;

"Cloning, characterization and overexpression of the phytase-encoding gene (phyA) of *Aspergillus niger*.";
 Gene 127:87-94(1993).
 [2]
 RA SEQUENCE FROM N.A.
 RA Mullane E.J.;
 RA Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE OF 24-464.
 RP STRAIN-NRRL 3135 / VAN TIEGHEM / FICUUM;
 RC MEDLINE=93249451; PubMed=8387289;
 RX Ullah A.H.J., Dischinger H.C. Jr.;
 RA "Aspergillus ficuum phytase: complete primary structure elucidation by chemical sequencing.";
 RT Biochem. Biophys. Res. Commun. 192:747-753(1993).
 [4]
 RN SEQUENCE OF 71-93.
 RP STRAIN-NRRL 3135 / VAN TIEGHEM / FICUUM;
 RC MEDLINE=91298982; PubMed=1648914;
 RX Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
 RA "Cyclohexanedione modification of arginine at the active site of Aspergillus ficuum phytase.";
 RT Biochem. Biophys. Res. Commun. 178:45-53(1991).
 [5]
 RN CHARACTERIZATION, AND PARTIAL SEQUENCE.
 RP STRAIN-NRRL 3135 / VAN TIEGHEM / FICUUM;
 RC MEDLINE=89160685; PubMed=2852807;
 RX Ullah A.H.J.;
 RA "Aspergillus ficuum phytase: partial primary structure, substrate selectivity, and kinetic characterization.";
 RT Prep. Biochem. 18:459-471(1988).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RP STRAIN-NRRL 3135 / VAN TIEGHEM / FICUUM;
 RC MEDLINE=97307250; PubMed=9164457;
 RX Kostrewa D., Gruninger-Leitch F., D'Arcy A., Broger C., Mitchell D., van Loon A.P.;
 RA "Crystal structure of phytase from *Aspergillus ficuum* at 2.5-A resolution.";
 RT Nat. Struct. Biol. 4:185-190(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER THE NAME PHYTASE NOVO.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z16414; CAA78904.1;
 CC EMBL; M94550; AAA32705.1;
 CC PIR; JN0482; JN0482.
 CC PIR; P00023; P00023.
 CC PIR; JN0656; JN0656.
 CC PDB; 1IHP; 18-MAR-98.
 CC InterPro; IPR000560;
 CC Pfam; PF00328; acid_phosphat_1;
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1;
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; 3D-structure.
 KW SIGNAL 1 23
 FT CHAIN 24 467 3-PHYTASE A.
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
 FT

Query Match	72.6%	Score 1684	DB 1	Length 467
Best Local Similarity	73.9%	Pred. No. 3.4e-132		
Matches 331	Conservative 40	Mismatches 59	Indels 18	Gaps 10
QY 1	NSHSCDITVD-GVQC-PELSHLWGQYSPFSLADESAISPDVPGKCRVTFVQVLSRHHGARY 58			
DB 27	NQSSCDITVDQGYCFSEYSHLWGQYAFPEFSLANESVSEVPACRCVTFQVLSRHGARY 86			
QY 59	PTSSKSKYSALIERIQNA-TFKGKYAFLKTYNYTLGADDLTFPGENQMVNSGIKPYRR 117			
DB 87	PTDSKGRKYSALIEEQNAITFDGKYAFLKTYNYSLGADDLTFPGEQLVNSGIKFYQR 146			
QY 118	YKALARNIVPVRASGSDRVIASKEKTEGFQSAKLADP---AHQASPVINVIPEGSY 174			
DB 147	YESLTRNIVPFISSGSRVIASSKKEIEGFQSTKLDPRAQPCQSSPKIDVISEASS 206			
QY 175	NNTLDHGLCTAFEDSTLGGDAEANTAFVAPPFARLE-ALPGVNLTDDEVNLMDCPF 233			
DB 207	NNTLDPGCTVFEDESELADIVAEANTATFVPSIRQRLENDLSGVTLTDTEVYLMDCSF 266			
QY 234	DTVARTSDATQLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQCVGF-NELIAR 290			
DB 267	DTISTSTVDTKLSPFCDLFTHDEWIINVDYLOSLKYGYGAGNPLGPTQGVYANELIAR 326			
QY 291	LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSFFALGLYNGTKPLSTTSV 350			
DB 327	LTHSPVHDDTSSNHTLSSPATFPLNSTLYADFSDHNGIISLFLALGLYNGTKPLSTTV 386			
QY 351	EST-ETDGYAASTVPFAARVAYEMQCCAGGGGEGEKEPELVRVLVNDRVVPLHGCGVD 409			
DB 387	ENITQDGFSSANTVPFASRLVYEMQCCA-----EQEPLVRVLVNDRVVPLHGCGVD 439			
QY 410	KLGRCKLDDFVEGLSFARSGGNNAECFA 437			
DB 440	ALGRCTRDSTFVRLSFARSGGDWNAECFA 467			
RESULT 3				
PHYB_EMENI	STANDARD;	PRT;	463 AA.	
ID	PHYB_EMENI			
AC	000093;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE			
DE	3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE			
DE	Phosphohydrolase B).			
OS	Emericella nidulans (Aspergillus nidulans).			
CC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
CC	Eurotiales; Trichocomaceae; Emericella.			
NCBI_TaxID	5072;			
UN	[1]			
RP	SEQUENCE FROM N.A.			
XP	MEDLINE=98007872; PubMed=9349716;			
IA	Paramontes L., Haiker M., Henriquez-Huacaa M., Mitchell D.B.			

119	KALARNIVPVRASGSDRVASAEKFI	EGFQSAKLAD - PAHQASPVINVTI	PEGSGYNNIT	177	
119				177	
146	KNLARKNTPIPRASGSDRVVASAEKFI	NGRKAQLHDHGSKRATPVVNVTI	PIIDGFGNNT	205	
146				205	
178	LDHGLCTAFADSTGLDDAEANFTAV	FAPPIRABLE - ALPGVNITD	EDVVNLMDMCPDFTV	236	
178				236	
206	LDHSTCVSFENDERADEIANFTAIM	GPPIRKLKLENDLPGILKTNNVTI	YLMDCSFDTM	265	
206				265	
237	ARTSDATQUSPFCDLFTADFW - QYDYLQSI -	KYGYGAGNPLGPAQGVGF - NELLARLTH	293		
237				293	
266	ARTAHGETISPFCAITKEWLOYDY	LQSLSKYGYGAGSPGPAQGI	GFTNELLARLTQ	325	
266				325	
294	SPVQDHTSTNHTLDSNPATFPLNATLYAD	FSDHNTMVSTIFFALGLYNGTKP	LSTTSVSESI	353	
294				353	
326	SPVQDNTSTNHTLDSNPATFPLDRKLYAD	FSDHNSMISIFFAMGLYNGTQ	PLSMDSVESI	385	
326				385	
354	-ETDGYAASWTVPFAARVYEMMOCEAGSGG	GEKPEPLRVLVNDRVVP	PLHGGGVDKLG	412	
354				412	
386	QEMDGYAASWTVPFCARAYFELMQCE -	-----	-KKPEPLRVLVNDRVVP	PLHGGGVDFKG	437
386				437	

	Query Match	15.3%;	Score 356;	DB 1;	Length 467;
	Best Local Similarity	26.9%;	Pred. No. 5.2e-22;		
	Matches 119;	Conservative 63;	Mismatches 190;	Indels 70;	Gaps 17;
QY	15	ETSHLWGQYSPFSLADESAISPDVDPKGRVTFVQVLSRHGARYPTSSKSKKYSALIERI	74		
Db	36	EIFPFLGGSGPYSPGDIGISRDLPESCEMKQVQVGRHGERYPVTSRAKSIIMTTWYKL	95		
QY	75	QK-NATFKCKYAPLK-----TNYNYLT-----GADDTLPF-GENQMVNSGKIFYRR	117		
Db	96	SNYTGQFSGALSFLNDYEFFIRDTKNLEMETTLANSVNVLPNYTGEMNAKRHARDFLAQ	155		
QY	118	YKALARNIVPF-VRAGSGDRVLTAAEKFEFGQSAKLADPAHQASPVINVIPIEGSGYNN	176		
Db	156	YGYMYENQTSFAVFTSNRCHDTAQYFIDG-----LGDKNFISLQTISE--AESAGANT	208		
QY	177	TLDHGLCTAFEDSTLGGDAEANFTAVFAPPRIARL-EALPGVNLTDDEVNLMWDMCPFD	235		
Db	209	LSAHHSCPAWDD-VNDDILKIKYDKTGLSGIAKLNKENKGLNLTSSDANTFFAWCAYEI	267		
QY	236	VARTSDATQLSFCDLFTADEW-QYDYIQLSLK-YGYGAGNPLGPAQGVG-FNELIARLT	292		
Db	268	NAR-----GYSDICNIFTKDELVRFSYGODLEYTYQTGPGYDVRVSGANLFNASVKLLK	322		
QY	293	HSPVODHTSTNHTLDSNPATFPLNATLYADFESHONTMVSIFFAFLGLYNGTKPLSTSIVES	352		
Db	323	ESEYQDQ-----KVWLSFTHDITDILNYLTITGIIIDQNNLTAEHVPPF	364		
Y	353	IETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLRVLVNDRVYPLHGCGVDKLG	412		
b	365	ME-NTFHRSWYVPOGARVYTEKFC-----SNDTYVRVINDAVVPIETCTSGPGF	414		
Y	413	RCKLDDF-----VEGLSFAR	427		

CONFLICT	219	DED -> MKT (IN REF. 1).
SEQUENCE	467 AA; 52776 MW; 05FBB80DEB41B0FF CRC64;	

Sat Oct 27 15:25:43 2001

us-09-488-265-27.rsp

```
Query Match      15.2%; Score 353; DB 1; Length 467;
Best Local Similarity 26.8%; Pred. No. 9.3e-22;
Matches 117; Conservative 62; Mismatches 187; Indels 70; Gaps 16;

QY 21 GQYSPFFSLADESATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIERIQK-NAT 79
DB 42 GGAGPYFFPGDYGISRLDPESCEMKQLMARHGERYPYTSKGTATIMKTWYKLSNTRQ 101
QY 80 FKGKYAFLEK-TYNYTLGADD-----LTPF-GENQMVNSGKIFRYRKALAR 123
DB 102 FNGSLFLNDDYEFFIRDDDDLEMETTFANSNDVNLNPNVTGEMDAKRAREFLAQYWFYE 161
QY 124 NIVPFF-VRASGSDRVIAAEAFIEGFQSAKLADPAHQASPVINVIPEGSGYNNNTLDHGL 182
DB 162 NOTSPFIFAAASSERVHDTAQFIDG-----LGDQFNISLQIVSEAMSAGA---NTLSAGN 213
QY 183 CTAFFEDSTLGDAAEFANFTAVFAPPIRAKL-BALPGVNLTTDEDVNLMDMCFDVTARTSD 241
DB 214 ACPGWDEDANDDILDKYDTTYLDDIAKRLNKENKGLNLTSKDANTLFAWCAYELNAR--- 270
QY 242 ATQLSPFCLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVG-FNELIARLTHSPVQD 298
DB 271 --GYSDVCDIFTEDELVRYSYGODLVFSFQDGGYDMIRSVCANLNFATLKLKQSETQD 328
QY 299 HTSTNHTLDSNPATFPLNATLYADFSDHNTMYSIFPALGLYNGTKPLSTSVESLETGQY 358
DB 329 -----LKVWLSFTHDTHDILNLTAGIIDDKNLNTAEYVPPFM-GNTF 369
QY 359 AASWTVPFAARAIVEMQCEAGGGEGEKEPLVRLVNLNDRVPLHGGCVDLKGRCKLDD 418
DB 370 HKSWMYVPOGARVYTFEQC-----SNDTYRVYVINDAVVPIETCTSGPGFSCIND 420
QY 419 F-----VEGLSFAR 427
DB 421 FDYAEKRVAGTDFLK 436

RESULT 7
ID PPAS_YEAST STANDARD; PRT; 467 AA.
AC P00635;
DT 21-JUL-1986 (Rel. 01; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
GN PHO5 OR YBR093C OR YBR0814.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-26.
RX MEDLINE=83168913; PubMed=6300772;
RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
of repressible acid phosphatase contains a signal peptide.";
RL Nucleic Acids Res. 11:1657-1672(1983).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-45.
RX MEDLINE=85037940; PubMed=6093051;
RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
RT "Structural analysis of the two tandemly repeated acid phosphatase
genes in yeast.";
RL Nucleic Acids Res. 12:7721-7739(1984).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95208357; PubMed=7900426;
RA Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
RL Yeast 10:1363-1381(1994).
RN [4]

RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=87064474; PubMed=3537710;
RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
RA Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
cluster within the acid phosphatase multigene family of Saccharomyces
cerevisiae.";
RL Mol. Cell. Biol. 6:1855-1865(1986).
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; V01320; CAA24630.1; -
CC EMBL; X01079; CAA25555.1; -
CC EMBL; X78993; CAA55598.1; -
CC EMBL; Z35962; CAA85046.1; -
CC EMBL; X01080; CAA25556.1; -
CC PIR; S05795; PABYC.
CC PIR; B25241; B25241.
CC HSP; P34752; IHP.
CC SGD; S0000297; PHO5.
CC InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
DR Hydrolase; Glycoprotein; Signal; Multigene family.
KW SIGNAL
FT CHAIN 1 17
FT SIGNAL 18 467
FT ACT_SITE 75 75
FT ACT_SITE 337 337
FT CARBOHYD 97 97
FT CARBOHYD 103 103
FT CARBOHYD 162 162
FT CARBOHYD 192 192
FT CARBOHYD 250 250
FT CARBOHYD 315 315
FT CARBOHYD 356 356
FT CARBOHYD 390 390
FT CARBOHYD 439 439
FT CARBOHYD 445 445
FT CARBOHYD 456 456
FT CARBOHYD 461 461
FT CONFLICT 36 36
FT CONFLICT 130 130
FT CONFLICT 294 294
FT CONFLICT 446 446
FT CONFLICT 462 462
FT CONFLICT 466 466
FT CONFLICT 467 467
SQ SEQUENCE 467 AA; 52858 MW; DC3C9504BC2D3D0C CRC64;

Query Match      15.0%; Score 347; DB 1; Length 467;
Best Local Similarity 26.5%; Pred. No. 2.9e-21;
Matches 117; Conservative 57; Mismatches 187; Indels 80; Gaps 18;

QY 21 GQYSPFFSLADESATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIERIQK-NAT 79
DB 42 GGAGPYFFPGDYGISRLDPESCEMKQLMARHGERYPYTSKGTATIMKTWYKLSNTRQ 101
QY 80 FKGKYAFLEK-TYNYTLGADD-----LTPF-GENQMVNSGKIFRYRKALAR 123
DB 102 FNGSLFLNDDYEFFIRDDDDLEMETTFANSNDVNLNPNVTGEMDAKRAREFLAQYWFYE 161
```

QY 124 NIVPF-VRASGSDRVIAAEKFIIEGFSQAKLADPAHQASPVINVLII-----PEGSGYNNT 177
 Db 162 NOTSFVFTSNKSRCHDTAQVFIIDG-----LGQD-----FNITLTVSEASAGANTL 209
 QY 178 LDHGLCTAFEDSLGDDAEANFTAVFAPPTARL-EALPGVNLTDVVMNMDMCPDPTV 236
 Db 210 SACNSCPAW-DYDANDDIVNEYDTTLLDDIAKRLNKENKGLNLTSTDAFLSWCAFEV- 267
 QY 237 ARTSDATQSPFCDFLTADSEW-QYDYLQSL-KYGYGAGNPLGPAQGVG-FNELIARLTH 293
 Db 268 ----NAKGYSDVCDITFKDELHYVSYQDLHTYHEGPGVDIILKSGVSNLNFNASVKLLKQ 323
 QY 294 SPVODHTSNHTLDSNPATFPLNATLYADFSDHTMTVSIFFAALGLYNGTKPLSTTTSVESI 353
 Db 324 SEIQDQ-----KWLSTHTDITLTLTTAGIIDDKNLNTAEVVPFM 365
 QY 354 ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDVRVPLHGGVDKGLGR 413
 Db 366 -GNTFHRSWVPGQARVYTEKFC-----SNDTYVRYVINDAVVPIETCTSGTGF 415
 QY 414 CKLDDE-----VEGLSPAR 427
 Db 416 CEINDFYAEKRVAGTDFLK 436

RESULT 8

PPAD_YEAST ID PPAD_YEAST STANDARD; PRT; 468 AA.
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE PROBABLE ACID PHOSPHATASE YDL024C PRECURSOR (EC 3.1.3.2).
 GN YDL024C OR D2815.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Vissers S., Urrestarazu L.;
 RL Submitted (Feb-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z48432; CAA88335.1; --
 DR EMBL; Z74072; CAA98583.1; --
 DR SGD; S0002182; YDL024C.
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid.phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 468
 FT ACT_SITE 76 76
 FT ACT_SITE 338 338
 FT CARBOHYD 98 98
 FT CARBOHYD 163 163
 FT CARBOHYD 193 193
 FT CARBOHYD 202 202
 FT CARBOHYD 238 238

FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDDF162C CRC64;
 Query Match 14.5%; Score 337; DB 1; Length 468;
 Best Local Similarity 26.0%; Pred. No. 2e-20;
 Matches 119; Conservative 60; Mismatches 161; Indels 118; Gaps 19;
 QY 21 GQVSPFSLADESAISPDVPGKGRVTFVQLSRHGRARYPTSSSKSKYSALIERLOKNTWF 80
 Db 43 GGSAPYFSPANYGIPTDIPEGCRLTQVQMIGRGERYPTSRSEAKD---IFEVYKISNY 99
 QY 81 KGKY----AFLKT-YNYTL-----GADDLTPF-GENQMVNSGIKIFYRKYA 120
 Db 100 TGYEGSLSLNNGYEFFIPDESLEMETTLQNSIDLNPYTGEMNAKRHAREFLAKYK 159
 QY 121 LARNIVPF-VRASGSDRVIAAEKFIIEGFSQAKLADPAHQASPVINVLPEGSGYNNTLD 179
 Db 160 LMENCTNPFITNSKRIYDTAQYFAEL-----GDGFNISLQ 197
 QY 180 HGLCTAFEDSLG-----DDAEANFTAVFAPPTARL-EALPGVNL 220
 Db 198 ----TLSENSSGANTLAAKSCPNWNSNANDILMSYSDYLENLSIDLNDENKGLNLS 253
 QY 221 DEDVYNLMDMCPDPTVARTSDATQSPFCDFLTADSEW-QYDYLQSL-KYGYGAGNPLGP 278
 Db 254 RKDAALFWSWCAFE-----NAKGYSNICDIFSAALHISYETDLTSTFYQNGPGYKLIK 308
 QY 279 AQGVG-FN---ELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHTMTVSIF 334
 Db 309 SIGANLNFATVKLIQSAH-----LDQKVLSTHTDITLNLIT 347
 QY 335 ALGYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRV 394
 Db 348 TAGLIDTNRNLTTNHV-PFRHSVHRSWYIPQGARVYTEKFC-----SNDSYVRY 397
 QY 395 LVNDRVPLHGGVDKGLGRCKLDDEVE-----GLSF 425
 Db 398 VYNDVAVPIESCSGPGFGCEGTFYAYKDLRGVSF 435

RESULT 9

PPAL_PICPA ID PPAL_PICPA STANDARD; PRT; 468 AA.
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ACID PHOSPHATASE PHOI PRECURSOR (EC 3.1.3.2).
 GN PHOI.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E., Gannon P.M., Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RL Characterization of the gene and its product.";
 RL Gene 163:19-26(1995).
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -1- INDUCTION: BY PHOSPHATE STARVATION.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U28658; AAA85503.1; --
HSSP; P34752; LIHP.
InterPro: IPR000560;
Pfam: PF00328; acid_phosphat; 1.
PROSITE: PS00616; HIS-ACID-POSPHAT_1; 1.
PROSITE: PS00778; HIS-ACID-POSPHAT_2; 1.
Hydrolase; Glycoprotein; Signal: Multigene family.
SIGNAL 1 22
CHAIN 23 468
FT ACT_SITE 84 84
FT ACT_SITE 345 345
FT CARBOHYD 163 163
FT CARBOHYD 196 196
FT CARBOHYD 256 256
FT CARBOHYD 321 321
FT CARBOHYD 360 360
FT CARBOHYD 453 453
SEQUENCE 468 AA; 52690 MW; AE555B8E27718C2C CRC64;

Query Match 14.4%; Score 334.5; DB 1; Length 468;
Best Local Similarity 27.0%; Pred. No. 3.2e-20;
Matches 124; Conservative 65; Mismatches 187; Indels 83; Gaps 22;

QY 7 TVDGYQCEISHLWGOYSPF-----FSLADESAIPDPKGRVTFVQVLSRHGARYPTS 61
DB 40 TDDQYNI--LRHL-GGLGYIGYNGWGTAAESEI-----ESTIDOAHLMLHGRYPT 91
QY 62 SKSKYSALIERLOKATFK--GKVAFLKTYNYTLG-----ADDLTP--FGNQMVNS 110
DB 92 NVGKQLEALYQKL-LDADVEVPTGPLSFDDYDYFVSDAAWYEQETKGFYSLNTAFDF 150
QY 111 GIKFYRYKALARNIYP-----FVRASGSRVIAASAKFTFEGFSQAKLADPAHQSPVINV 166
DB 151 GTTLREYDHLINTSEGGKLSWAGSQRVVDTKYFAQGMKSNYTD-----MVEV 203
QY 167 IIEP--GSCYNTLDHGLCTAFEDSTGLDGAENATVAFAPPFIRARLEAL-PGVNLIDE 222
DB 204 VALEEKSGQLSUTARISCPNYSNHIYD--GDFPNDAIEAREADRLNTLSPGNITAD 260
QY 223 DVNLMDCPFDTVARTSDATQLSPFCDLPTADEWYD-YLQSLK-YGYGACNPLGPAQ 280
DB 261 DIPTIALYCGFELNVRCG-----SSFCVLSREALLYTAYRLDGLWYNNVNGNPLGKTI 315
QY 281 GVGFNELIARLTHSPVDQHTSTNHTLDSNPATPLNATLYADESHDNTWVSIPFALGLYN 340
DB 316 GYVYANATROLLEN-----TEADPRDP-----LYFSFSDTDLLOVFTSLGLFN 360
QY 341 GTK-PLS-----TTSVESIETDGYAASVTVPFAARVYVMQCEAGGGGGEKEPLVRVL 395
DB 361 VTDLPDQIQOTSEKSTE-----IVPMGARLLRLRLCTV-----EGEEKYIVRTI 407
QY 396 VNRVVPVHGGCVKDLGRCKLDDFVEGLSFARSGGNWAE 434
DB 408 LNDVAFPLSDCSGPGFSCPLNDVYSLREALNEDSDFAE 446

RESULT 10
PHYB_ASAPW STANDARD; PRT; 479 AA.
AC PHYB_ASAPW
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE)
GN PHYB OR APH.

OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ALK0243;
RA MEDLINE=94040796; PubMed=8224894;
RA Piddington C.S., Houston C.S., Palohimo M.T., Cantrell M.A.,
RA Miettinen-Oinonen A., Nevalainen H., Rambosek J.A.;
RT "The cloning and sequencing of the genes encoding phytase (phy) and
RT pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
RT awamori.";
RL Gene 133:55-62(1993).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=99264417; PubMed=10329192;
RX Kostrewa D., Wyss M., D'Arcy A., van Loon A.P.;
RT "Crystal structure of Aspergillus niger pH 2.5 acid phosphatase at
RT 2.4-A resolution";
RL J. Mol. Biol. 288:965-974(1999).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: L02420; AAA16897.1; --
PIR: JN0890;
PIR: JN0890; JN0890.
PDB: LOFX; 19-APR-00.
InterPro: IPR000560;
Pfam: PF00328; acid_phosphat; 1.
PROSITE: PS00616; HIS-ACID-POSPHAT_1; 1.
PROSITE: PS00778; HIS-ACID-POSPHAT_2; 1.
Hydrolase; Glycoprotein; Signal: 3D-structure.
SIGNAL 1 19
CHAIN 20 479
FT ACT_SITE 82 82
FT ACT_SITE 337 337
FT DISULFID 71 387
FT DISULFID 128 472
FT DISULFID 216 441
FT DISULFID 225 298
FT DISULFID 413 421
FT CARBOHYD 191 191
FT CARBOHYD 315 315
FT CARBOHYD 458 458
SEQUENCE 479 AA; 52678 MW; 4F8E0F3778CC3B08 CRC64;

Query Match 13.2%; Score 307.5; DB 1; Length 479;
Best Local Similarity 27.4%; Pred. No. 5.7e-18;
Matches 124; Conservative 55; Mismatches 175; Indels 99; Gaps 23;

QY 9 DGYQCEISHLWQ--YSPFFSLADESAIPDPKGRVTFVQVLSRHGARYPTSRSKSK 66
DB 41 DGYSI--LKHGGNGPYSERVSY-----GIADPPTSCVDQVIMVKRHGERYPSPSACK 94
QY 67 YSALIERIQK-NAT-FKGYAFLKTYNYTL-----GADDLT-PF-GENQMVNSIKFYR 116
DB 95 IEEALAKVYSINTTEYKGLDLAEFLNDWTYVVPNECYNAETTSCTPYAGLLDAYNHGNDYKA 154
QY 117 RYKAL--ARNIVPFVASCSDRVIAAEKFIQSGFQSAKLADPAHQASPVINVIPEGSY 174
PHYB OR APH.

Db 155 RYGLHNGETVVPFF--SSGYGRVETARKFGEF-----FGY 190

QY 175 NNTLDHGLCTAFEDSTLGDAAE-----NFTAVFAPPPIR---ARLEAL-PGV 217

Db 191 NYSNAALNIISEVAGADSLPTCTDNDQTTCDNLTYQL-PQFVAAARLNSQNGM 249

QY 218 NLTDDEVVNLMDMCPFTVARTSDATQSPFCDLFTADEW-QYDYLOSLKYVYGAG--- 273

Db 250 NLTASDYNILMVASFELNAR-----PFSNWINAFQDEWVSFGYEDLNYI-YCAGPGD 303

QY 274 NPLGPAQGVGFENELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMVSIF 333

Db 304 KMAAAGVAYANASITLLNQGP-----AGSLFFNFAHDTNITPIL 345

QY 334 FALGLY--NGTKPLSTTSVESIETDGYAASVTPFAARAVENMOCFAAGGGGEGEKEPL 391

Db 346 AALGVILPNEDELPLDRVAF-----GNPYSIGNIVPMGGHLTIERLSQAALSDEG---TY 398

QY 392 VRLVNDRVVPLHGGVDKLRCKLDDFVEGLS 424

Db 399 VRLVNEAVLPFNDCSTSGPGYCLANYTSLN 431

RESULT 11

ID PHYB ASPNG STANDARD; PRT; 479 AA.

AC P34754;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE

DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE

GN PHYB.

OS Aspergillus niger.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OC NCBI_taxid=5061;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.

RX MEDLINE=93371452; PubMed=7916610;

RA Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,

RA Ullah A.H.J.

RT "Identification and cloning of a second phytase gene (phyB) from

RL Biochem. Biophys. Res. Commun. 195:53-57(1993).

CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE

CC FROM PHYTATE.

CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =

CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L20567; AAA02934.1; -

CC InterPro: IPR000560; -

DR Pfam: PF00328; acid_phosphat; 1.

DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.

DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.

KW Hydrolase; Glycoprotein; Signal.

FT SIGNAL 1 19

FT CHAIN 20 479

FT ACT_SITE 81 81

FT ACT_SITE 82 82

FT ACT_SITE 337 337

FT ACT_SITE 106 106

FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 479 AA; 52611 MW; 395D4DA2B50DFC4 CRC64;

Query Match 13.2%; Score 305.5; DB 1; Length 479;

Best Local Similarity 27.4%; Pred. No. 8.4e-18;

Matches 124; Conservative 57; Mismatches 173; Indels 99; Gaps 24;

QY 9 DGYQCEPISHLWQ---YSPFESLADESAISPDPKCRVTFVQVLSRHGARYTSSKSK 66

Db 41 DGYSI--LKHYGNGPYSERVSY---GIARDPTCEVDQVIMVKRIGERYSPSAGKS 94

QY 67 YSALIERIOK-NAT-FKGYAFILKTYNYTL-----GADDLT-PF-GENOMVNSGIKEYR 116

Db 95 IEALAKVYSINTTEYKGLAFLNDWTYYVPNECYNAETTSQPYAGLLDAYNHGNDYKA 154

QY 117 RYKAL--ARNIVPEVRASGSDRVIASAEKIEGFQSAKLADPAHQASPVINVIPEGSGY 174

Db 155 RYGLHNGETVVPFF--SSGYGRVETARKFGEF-----FGY 190

QY 175 NNTLDHGLCTAFEDSTLGDAAE-----NFTAVFAPPPIR---ARLEAL-PGV 217

Db 191 NYSNAALNIISEVAGADSLPTCTDNDQTTCDNLTYQL-PQFVAAARLNSQNGM 249

QY 218 NLTDDEVVNLMDMCPFTVARTSDATQSPFCDLFTADEW-QYDYLOSLKYVYGAG--- 273

Db 250 NLTASDYNILMVASFELNAR-----PFSNWINAFQDEWVSFGYEDLNYI-YCAGPGD 303

QY 274 NPLGPAQGVGFENELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMVSIF 333

Db 304 KMAAAGVAYANASITLLNQGP-----KEAGP-----LFFNFAHDTNITPIL 345

QY 334 FALGLY--NGTKPLSTTSVESIETDGYAASVTPFAARAVENMOCFAAGGGGEGEKEPL 391

Db 346 AALGVILPNEDELPLDRVAF-----GNPYSIGNIVPMGGHLTIERLSQA---TALSDKGTY 398

QY 392 VRLVNDRVVPLHGGVDKLRCKLDDFVEGLS 424

Db 399 VRLVNEAVLPFNDCSTSGPGYCLANYTSLN 431

RESULT 12

ID PPA5_KLULA STANDARD; PRT; 469 AA.

AC P52289;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).

GN PHO5.

OS Kluyveromyces lactis (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

OC NCBI_taxid=28985;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CBS 2359/152;

RA Ferminan E.

RL Thesis (1995), University of Salamanca, Spain.

CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN

CC ALCOHOL + ORTHOPHOSPHATE.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).

CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 233995; CAA83964.1; --
InterPro; IPR000560; --
DR Pfam; PF00328; acid-phosphat; 1.
DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 16
FT CHAIN 17 469
FT ACT_SITE 77 77
FT ACT_SITE 339 339
FT CARBOHYD 23 23
FT CARBOHYD 31 31
FT CARBOHYD 32 32
FT CARBOHYD 129 129
FT CARBOHYD 201 201
FT CARBOHYD 229 229
FT CARBOHYD 250 250
FT CARBOHYD 317 317
FT CARBOHYD 392 392
FT CARBOHYD 447 447
FT SEQUENCE 469 AA; 52532 MW; 863B528D0740AA7E CRC64;

Query Match 12.9%; Score 300.5; DB 1; Length 469;
Best Local Similarity 25.9%; Pred. No. 2.1e-17; Indels 99; Gaps 21;
Matches 118; Conservative 66; Mismatches 17;

QY 17 SHLWGOYSPFSLADESATSPDPKGRVTFVQLSRHGARYPTSSKSKYSALIERIQK 76
DB 41 SLNGQ-GPHYDPSFGFIPVEPDQCTVEHQMRLARHGRYPTASKGKLMIALMDKLKE 99
QY 77 NATFKGYAFKTYNTLGGADLTTPGEMQVNSGKIFRYR----- 118
DB 100 ---FQGGY-----NDPLEVFNDEFFVSNTPKYFDQLTNSTVDPSNPYAGAKTA 145
QY 119 ---KALARNIVPVRASGSDRVIAAEKFIQFQSAKLADPAHQASPVINV---IIPEG 171
DB 146 QHLGKIYAYNGDLF--SDSNPFTSSGRVH--QTAKYVSSLEELDIQLDLQIQEN 201
QY 172 --SGYNTLDHGLCTAFEDSTLGDGAENFTAVFAPPIRAR--LEALPGVNL--DEDVYN 226
DB 202 ETSANSLTPADSCMTY--NGDLGDEYFENATPLPLTDIKRNMKNSNLTLEHDDIEL 260
QY 227 LMDMCPPTVARTSDATQLSPFCDLFTADEW--OYDYLQSL--KYGYGAGNPL--GPAQGVG 283
DB 261 LVDWCAFEETNKGSSAV-----CDLFRNDLVAYSYANNVYRAGAGNPMSPIGSVL 315
QY 284 FNELIARLTHSPVDHSTNHTLDSNPATFPLNATLYADFSDMTWVSIFALGLY--NGT 342
DB 316 VNASYNLLTQADELN-----KWLSESHDITQOQFISALGLIDNGV 357
QY 343 KPLSTTSV--ESITDGYASWTVPPFAARAYVENMQCEAGGGEKEPLVRVLVNDRV 400
DB 358 TEYSLDQVDFQNIQ-----QLSWFTPMGRIFTEKLK-----GNASYVRYINDVI 404
QY 401 VPLHCGGVGDKLGRCKLDDF-----VEGLSFARS 428
DB 405 IPVPCCTSGPFCSEPIEDFDYIINRLNGIDYVSS 439

RESULT 13
ID PPAL SCHPO STANDARD; PRT; 453 AA.
AC P08091.
DT 01-AUG-1988 (rel. 08, Created)
DT 01-AUG-1988 (rel. 08, Last sequence update)

DT 01-OCT-2000 (rel. 40, Last annotation update)
DE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
GN PHO1 OR SPBP4G3.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86140050; PubMed=3005272;
RA Elliott S., Chang C., Schweingruber M.E., Schaller J., Rickli E.E.,
RA Carbon J.;
RT "Isolation and characterization of the structural gene for secreted acid phosphatase from Schizosaccharomyces pombe";
J. Biol. Chem. 261:2936-2941(1986).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC Rieger M., McDougall R.C., Rajandream M.A., Barrell B.G.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CELL WALL.
CC -!- INDUCTION: RERESSED BY PHOSPHATE AND WEAKLY BY THIAMINE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M11857; AAA35321.1; --
EMBL; AL137099; CAB68657.1; --
PIR; A25326; A25326.
DR HSP; P34752; IHP.
DR InterPro; IPR000560; --
DR Pfam; PF00328; acid-phosphat; 1.
DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Cell wall; Signal.
FT SIGNAL 1 18
FT CHAIN 19 453
FT ACT_SITE 69 69
FT ACT_SITE 330 330
FT CARBOHYD 95 95
FT CARBOHYD 151 151
FT CARBOHYD 183 183
FT CARBOHYD 193 193
FT CARBOHYD 243 243
FT CARBOHYD 319 319
FT CARBOHYD 410 410
FT CARBOHYD 429 429
FT CARBOHYD 443 443
FT SEQUENCE 453 AA; 50557 MW; 7CF891256EB154D1 CRC64;
ACID PHOSPHATASE.
NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
PROTON DONOR (BY SIMILARITY).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 12.8%; Score 298; DB 1; Length 453;
Best Local Similarity 25.6%; Pred. No. 3.2e-17;
Matches 107; Conservative 63; Mismatches 182; Indels 66; Gaps 16;

QY 23 YSPFFSLADESATSPDPKGRVTFVQLSRHGARYPTSSKSKYSALIERIQK---NAT 79
DB 43 HKPYF-----YGPISDFPTTKIKQVHTLQHRGSRNPTGNGNAFADAVGLANFQORLLNGS 97
QY 80 FKGY-----AFLKTYNTL---GADLTTPGEMQVNSGKIFRYRKYKALARNIVPV 129
DB 98 VPIDYVSGNPLSFVPTWTFVIEANADALSSSGRVLEFDMGRQFYRHELFNASTYNI 157
QY 130 RASGSDRVIAAEKFIQFQSAKLADPAHQASPVINVIPEGSYNTLDHGLCTAFEDS 189

Db 158 YTAQQRVDSALWYGYCM-----FGEDVHNFTNVLSENATAGSNLSUSSNACPASDAD 213
 QY 190 TLGDAAENFTAVFAPTRARLEA-LPGVNLTDDEVNLMDCMDFDFTVARTSDATQLSPF 248
 Db 214 DFTTAPALEARNVTPPIRQLNPNYFSNLTNDIILNLYGICSYETALQ-----DYSEF 268
 QY 249 CDLFTADEWQYDYLQSLKYGYGAGNPL--GPAQGVGENELIARLTHSPVQDHTSTNHT 305
 Db 269 CKLFNSVDLFNFEYEGDLS-FSYGMGNSVKWGSIFGAYANSIANSURS-VEENNTQ----- 322
 QY 306 LDSNPATFPLNATLYADFSHDNTMWSIFFALGYNGT---KPLSTTSVESIETDGYAASW 362
 Db 323 -----QVFTAFTHDANIIPVETALGFDTONTPENPLPTS--YQVHSHMKASE 368
 QY 363 TVPFAARAYVEMMQCAGGGGEGEKEPLVRVLVNDVRVPLHCGGVYDKL-----GRCKL 416
 Db 369 FVPFAGNLITELPOCE-----DSKYVVRHLVNEEVPLSDCGFGPSNTSDGMCEL 418

RESULT 14
 PPA2_SCHPO STANDARD; PRT; 463 AA.
 AC Q01682;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 GN PH04 OR SPBC428.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
 RX MEDLINE=91064763; PubMed=2249257;
 RA Yang J., Schweingruber M.E.; for thiamin-repressible acid phosphatase
 RT "The structural gene coding for thiamin-repressible acid phosphatase
 RL in Schizosaccharomycetes pombe.";
 RN Curr. Genet. 18:269-272(1990).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- INDUCTION: REPRESSED BY THIAMINE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56939; CAA04258.1; -
 DR EMBL; AL034382; CAA22278.1; -
 DR PIR; S14119; S14119.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT.1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT.2; 1.
 KW Hydrolase; Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 463 THIAMINE-REPRESSIBLE ACID PHOSPHATASE.
 FT ACT_SITE 69 69 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 340 340 PROTON DONOR (BY SIMILARITY).

FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 463 AA; 52118 MW; F48EAFB8B6B234A CRC64;

Query Match 12.5%; Score 290; DB 1; Length 463;
 Best Local Similarity 24.3%; Pred. No. 1.5e-16;
 Matches 109; Conservative 70; Mismatches 183; Indels 86; Gaps 18;

QY 23 YSPFESLADSSALSPDPKCRVTFVQVLSRHGARYPTSS-----KSKYKALIERIQK 76
 Db 43 HEPYF-----NGPTTFPESCAIKQVHLLQHGSRNPTGDDTATDVSSAQYIDIFQNKLL 97
 QY 77 NATFFKGKYA-----FLKTYNYTL--GADDLTPGENOMVNSGIKFFRYRYKALARNIV 126
 Db 98 NGSIPTVNESYPENPLYFVKHWTPIKAENADQLSSSGRIELFDLGRQVFFRYEYELFDTDV 157
 QY 127 PFRVRSGRDVIASAEKFTIEGFSQAKLADPAHQASPVINVIPE--GSGYNNTLDHGLCT 184
 Db 158 YDINTAAQERVVDSEAEWFSYGMFGDDMQNKTN-----FIVLPEDDSAGANSIAMYYSCP 211
 QY 185 APEDSTLGDG---AEANFTAVFAPPFIRARLEAL--PGVNLTDDEVNLMDCMPPFTVAR 238
 Db 212 VYEDNNIDENTEAHTSWRNVLFLPIANLNKYFDSGYNLTVSDVRSLYICVYETIAR 271
 QY 239 TSDATQLSPFCDLFTADEW--QYDYLQSLKYGYGAGNPLGPAO-----GVCFNELIAR 290
 Db 272 DN-----SDFCSLFTPSEFLNEFYDSDLDYAWG-----GPASEWASTLCGAYVNNLANN 321
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMWSIFFALGLY---NGTKPLST 347
 Db 322 L-----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDITPEHPPLT 365
 QY 348 TSVESIETDGYAASWTVPFAARAYVEMMQCAGGGGEGEKEPLVRVLVNDVRVPLHCGC 407
 Db 366 D--KNIFTYSLKTSFSFVPFAGNLITELFC-----SDNKYYVVRHLVNOQVYPLTDCG 415
 QY 408 VDKLGR-----CKLDDFVEGLSFAFSGGN 431
 Db 416 YGPSGASDGLCELSAYLNSVVRVNSTN 443

RESULT 15
 PPAX_CABEL STANDARD; PRT; 755 AA.
 ID PPAX_CABEL
 AC Q09549; O17843;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE ACID PHOSPHATASE F26C11.1 (EC 3.1.3.2).
 GN F26C11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Matthews P., Lloyd C.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: October 26, 2001, 15:17:24 ; Search time 75.85 Seconds
(without alignments)
762.258 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDTVDGQCPEISHLW.....DFVEGLSFARSGGNWAECA 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 133205027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mbc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.unclassified:*
- 13: sp.vertebrate:*
- 14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1762.5	75.9	465	3	O00092
2	1683	72.5	467	3	Q9U0Z7
3	1672	72.0	467	3	Q93838
4	1662	71.6	467	3	Q9HQ00
5	1596.5	68.8	466	3	O00096
6	1588	68.4	466	3	O00100
7	1562.5	67.3	466	3	O00085
8	1275.5	55.0	487	3	O00107
9	317.5	13.7	442	3	Q74677
10	298.5	12.9	469	3	Q9Y846
11	285	12.3	463	3	O60172
12	207.5	8.9	460	5	Q9VV72
13	196	8.4	467	5	O96421
14	192.5	8.3	453	5	O96420
15	191.5	8.3	453	5	Q9W438
16	176	7.6	487	4	Q9UGA3
17	175	7.5	451	11	O35217
18	174	7.5	487	4	O9UNW1
19	173	7.5	487	4	O95172

20	172	7.4	481	11	Q9Z2L6
21	166	7.2	449	13	Q92170
22	154	6.6	468	10	O04509
23	142	6.1	198	3	Q9UTX1
24	133	5.7	274	11	Q9JUD5
25	129.5	5.6	381	11	Q9JGD5
26	128.5	5.5	374	11	Q9JMG5
27	115	5.0	683	5	O00838
28	113	5.0	707	5	O00839
29	111	4.8	452	5	O19175
30	109.5	4.7	447	5	O19175
31	106.5	4.6	447	5	O9U5S9
32	106.5	4.6	447	5	O9U5T5
33	106.5	4.6	447	5	O9U5S8
34	106.5	4.6	447	5	O9TW19
35	106	4.6	447	5	O9TW17
36	105.5	4.5	2205	5	O16857
37	105.5	4.5	447	5	O97187
38	105.5	4.5	447	5	O9U5T4
39	105.5	4.5	447	5	O9U5T0
40	105	4.5	447	5	O9TW40
41	105	4.5	537	5	O25332
42	104.5	4.5	888	5	O25336
43	104.5	4.5	447	5	O9U5V1
44	104.5	4.5	447	5	O9U5U7
45	104.5	4.5	447	5	O9U5U3

ALIGNMENTS

RESULT 1
O00092
ID O00092
AC O00092; PRELIMINARY; PRT; 465 AA.

DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (BC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHVA.

OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_taxid=5085;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN=ATCC 34625;
RX MEDLINE=97288063; PubMed=9143104;

RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
RT phytase from the fungus Aspergillus fumigatus.";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).

CC -|- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -|- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

CC -|- SUBCELLULAR LOCATION: SECRETED.
CC -|- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMIC ACTIVITY.

CC -|- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; U59804; AAB96872.1; -;
DR HSSP; P34752; 1HP.

DR InterPro; IPR000560; -;

DR Pfam; PF00328; acid phosphat; 1

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1;

DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

KW Hydrolase; Glycoprotein; Signal.

FT SIGNAL 1 26

Sat Oct 27 15:25:44 2001

us-09-488-265-27.rspt

```

FT CHAIN 27 465 3-PHYTASE A. 1762.5; DB 3; Length 465;
FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
FT SIMILARITY).
FT 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 82 82
FT 359 359 BY SIMILARITY.
FT DISULFID 30 39 BY SIMILARITY.
FT DISULFID 70 412 BY SIMILARITY.
FT DISULFID 213 463 BY SIMILARITY.
FT DISULFID 262 280 BY SIMILARITY.
FT DISULFID 434 442 BY SIMILARITY.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 465 AA; 50836 MW; 86FCID9058C9B2C9 CRC64;

Query Match 75.9%; Score 1762.5; DB 3; Length 465;
Best Local Similarity 77.4%; Pred. No. 7.7e-136;
Matches 345; Conservative 36; Mismatches 48; Indels 17; Gaps 10;

QY 2 SHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYP 59
DB 27 SKSCDVTVDLGYQCSPATSHLWGOYSPFFSLEDELSVSSKLPKDCRITLVQVLSRHGARYP 86
QY 60 TSSKSKKYSALIERIOKNA-TFKGYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 118
DB 87 TSSKSKKYSALIERIOKNA-TFKGYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 146
QY 119 KALARNIVPVRASGSDRVIAAEKFIQFQSAKLADP--AHQASPVINVIIPGSGYNN 176
DB 147 KALARSVPPIRASGSDRVIAAEKFIQFQSAKLADP--AHQASPVINVIIPGSGYNN 206
QY 177 TLHGICLCTAFEDSTLGDAAEANTFAPPIRARELE-ALPGVNLTDDEVVNLMDMCPFT 235
DB 207 TLHGICLCTAFEDSTLGDAAEANTFAPPIRARELE-ALPGVNLTDDEVVNLMDMCPFT 266
QY 236 VARTSDATQSPFCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPAQGVF-NELIARLT 292
DB 267 VARTSDATQSPFCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPAQGVF-NELIARLT 326
QY 293 HSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFLALGLYNGTKPLSTTSV 352
DB 327 RSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFLALGLYNGTKPLSTTSV 386
QY 353 I-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVDKL 411
DB 387 AKELDGYASWVFPFARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVDKL 439
QY 412 GRCKLDDEVEGLSFARSGGNWAECPA 437
DB 440 GRCKLDDEVEGLSFARSGGNWAECPA 465

```

```

RESULT 2
Q90027 PRELIMINARY; PRT; 467 AA.
AC Q90027
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (EC 3.1.3.8).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RA Hongning W., Qi W., Jing X.;
RT "PCR, cloning and characterization of the phytase (phyA) gene of

```

```

RT Aspergillus niger (China Strain).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218813; AAF25481.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Signal; Lyase; Hydrolase.
FT SIGNAL 1 19 POTENTIAL.
FT SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 72.5%; Score 1683; DB 3; Length 467;
Best Local Similarity 73.7%; Pred. No. 2.4e-129;
Matches 330; Conservative 43; Mismatches 57; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
DB 27 NSHSCDVTVDGYQCFSETSHLWGOYAFPTFLANESAISSPDVPGKCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKKYSALIERIOKNA-TFKGYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 117
DB 87 PTSSKSKKYSALIERIOKNA-TFKGYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 146
QY 118 YKALARNIVPVRASGSDRVIAAEKFIQFQSAKLADP--AHQASPVINVIIPGSGY 174
DB 147 YESLTRNIPIRSGSSSRVIAAGKFIQFQSTKLDKPRAGQSPKIDVYVISEASS 206
QY 175 NNTDLGLCTAFEDSTLGDAAEANTFAPPIRARELE-ALPGVNLTDDEVVNLMDMCP 233
DB 207 NNTDLGLCTAFEDSTLGDAAEANTFAPPIRARELE-ALPGVNLTDDEVVNLMDMCP 266
QY 234 DVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPAQGVF-NELIAR 290
DB 267 DTISTSTVDTKLSFPFCDLFTDEWINYDYLQSLKYYGYGAGNPLGPTQGVYANELLAR 326
QY 291 LTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFLALGLYNGTKPLSTTSV 350
DB 327 LTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFLALGLYNGTKPLSTTSV 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVD 409
DB 387 QITDITDGFSSANTVFPFARLYVEMMOCEA-----EQEPLRVLVNDRVVPVPLHGCPVD 439
QY 410 KLGRCKLDDEVEGLSFARSGGNWAECPA 437
DB 440 ALGRCTRDTSFVKGLSFARSGGNWAECPA 467

RESULT 3
O93838 PRELIMINARY; PRT; 467 AA.
ID O93838
AC O93838;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PHYTASE.
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-57;
RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
RT "Phytase having high-affinity for phytic acid.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022700; CAB19824.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.

```

```
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543E
```

Query Match	72.0%	Score 1672;	DB 3;	Length 467;
Best Local Similarity	72.8%;	Pred. No. 1.9e-128;		
Matches 326;	Conservative 46;	Mismatches 58;	Indels 18;	Gaps

QY	1	NSHSCD	TVD-GYQC-PEISHLWQGYSPFTSLADESALS	PDVPKGCVRVTFVOVLSRHGARY	58
Db	27	NQSTCTD	VQGYOCFSETHLWQYAPFSLANKSALS	PDVPAGCHVTFAQVLSRHGARY	86
QY	59	PTSSKSKY	SALIERLOKNA-TFKGYAFELKTYNYTLGADDL	TPFGENMVNSGKIFYRR	117
Db	87	PTDSKSKY	SALIEEQONATTFEGKYAFELKTYNSLGADDL	TPFGEQELVNSGVKFYQR	146
QY	118	YKALARNI	VPFVRASGSDRVIASAEKFIEGFOSAKLADP--	-AHOASPVINVLIPGSGY	174
Db	147	YESLTRNI	VPFTRSSGSDRVIASGNKFIEGFOSKLDKDP	PAQGGSGKIDVWSEASTS	206
QY	175	NNTLHGLCTA	FEEDSTLGDADAANTAVAPIPARLE-ALF	GVNLTDDEVVNLMDMCPF	233
Db	207	NNTLDPGCTV	FEEDSELADIEANFTATFVPIRORLENDL	SGVSLTDFTVYLLMDCSF	266
QY	234	DTVARTSDATOL	SPCEDFTADGW-QYDYLQSL-KYGYGAGNPLG	PAGQVGF-NELIAR	290
Db	267	DTISTSTDVT	KLSPCEDLTHEEWLNIDYQLSNIYYGHGAGN	PLGPTQGVGYANELIAR	326
QY	291	LTHSPVQDHT	STNNHTLDSNPATFPINATLYADFSDHNT	MVSIFFALGLYNGTKPLSTTSP	350
Db	327	LTHSPVHDOT	STNNHTLDSNPATFPINSTLYADFSDHNGI	ISIIFALGLYNGTKPLSSITA	386
QY	351	ESI-ETDGYAAS	WTVPFAARAYVEMMOCEAGSGGEGEKEPL	VRVLVNDVRVPLHGCGVD	409
Db	387	ENITQTDG	FSSANITVPFASRMVEMMQQS-----	-EQEPLRVVLVNDVRVPLHGCPVD	439
QY	410	KLGRCKLDD	DFEGLSFA	RSGGNWAECEFA	437
Db	440	ALGRCTRDS	FFVKGLSFA	RSGGDWGECEFA	467

RESULT 4
19HEQ0 PRELIMINARY; PRT; 467 AA.
Q9HEQ0;
Q9HEQ0; (TEMBLrel. 16, Created)
01-MAR-2001 (TEMBLrel. 16, Last sequence update)
01-MAR-2001 (TEMBLrel. 16, Last sequence update)
01-MAR-2001 (TEMBLrel. 16, Last annotation update)
PHYTASE.
Aspergillus ficum.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5058;
[1]
SEQUENCE FROM N.A.
A Zhang L., An L., Wang Y., Yuan X.;
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
L R EMBL: AY013315; ANG40885.1; -
O SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match	71.68;	Score 1662;	DB 3;	Length 467;
Best Local Similarity	72.88;	Pred. NO. 1.3e-137;		
Matches 326;	Conservative 46;	Mismatches 58;	Indels 18;	Gaps 10;

1 N\$HSCOTVD-GYQC-PEISHLWGY\$FFFLADE\$AISDPV\$KGCRTV\$FVQVLSRHGARY 58
27 NOSTCOTVDQGYC\$F\$F\$H\$H\$W\$G\$Y\$A\$F\$F\$S\$L\$A\$N\$K\$A\$S\$I\$P\$D\$V\$A\$G\$C\$H\$T\$T\$A\$Q\$V\$L\$S\$R\$H\$G\$A\$R\$Y 86
59 PT\$S\$K\$K\$Y\$S\$A\$L\$I\$E\$R\$I\$O\$K\$N\$A-\$T\$F\$K\$G\$Y\$A\$F\$U\$K\$T\$N\$Y\$T\$I\$G\$A\$D\$D\$T\$P\$F\$G\$E\$M\$O\$M\$V\$N\$S\$G\$K\$F\$Y\$R 117
87 PT\$D\$S\$K\$K\$Y\$S\$A\$L\$I\$E\$I\$O\$O\$N\$A\$T\$T\$E\$G\$Y\$A\$F\$U\$K\$T\$N\$Y\$S\$L\$G\$A\$D\$D\$T\$P\$F\$G\$E\$O\$E\$L\$V\$N\$S\$G\$Y\$K\$F\$Y\$Q\$R 146

QY	118	YKALARNIVPVRASGSDRVIAEAEKFIEGFQSAKLADP---AHOASPVINVLIPESGY	174
Db	147	YESLTRNIVPIRSSGSRVIAISGNFIEGFQSTKUDPRAQGSSPKIDVVISEASTS	206
QY	175	NNTLHGLCTAFEDSTLGGDAEANFTAVFAPIRLARLE-ALPGVNLTDEDDVNLMDMCPF	233
Db	207	NNTLDPCTCTVFEDSELADIEANFTATVPFSLRQRENDLSGVSLTDEVYILMDMCSF	266
QY	234	DTVARTSDATQLSPFCDLTFADEW-OYDYLOSL-KYYGYAGNPGLPAQGVGF-NELIAR	290
Db	267	DTLSTSTVDTKUSPFCDLTFHEWINDYLQSLNKYIGHGAGNPLCPTOGVGYNELIAR	326
QY	291	LTHSPVQDHTSTNHTLDNSPNATFLNATLYADFSDHNTMVSTFFALGLYNGTKPLSTTSV	350
Db	327	LTHSPVHDOTSSNHTLDNSPNATFLNSTLYADFSDHNGIIISLFALGLYNGTKPLSSTA	386
QY	351	EST-EFDGVAASWTVPFAARAYVENMQCRAGGGGEKEPLEVRVLVNDVRVPLHGCGVD	409
Db	387	ENITQTDGESSARTVPFASRMVVENMQQS-----EQEPLRVLVNDVRVPLHGCGPD	439
QY	410	KLGRCKLDDFVEGLSFARSGGNWAECEFA	437
Db	440	ALGRCTRDSFVKGLSFARSGGDWAECEFA	467
RESULT	5		
ID	O00096	PRELIMINARY;	PRT; 466 AA.

RESULT	5	
000096		
ID	000096	PRELIMINARY; PRT; 466 AA.
AC	000096;	
DT	01-JUL-1997 (Tremblrel. 04, Created)	
DT	01-JUL-1997 (Tremblrel. 04, Last sequence update)	
DT	01-MAR-2001 (Last annotation update)	
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3- PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE PHYA).	
DE	PHOSPHOHYDROLASE A).	
GN	Talaromyces thermophilus.	
OS	Talaromyces thermophilus.	
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;	
OC	Euriales; Trichocomaceae; Talaromyces.	
OX	NCBI_TaxID=28565;	
RP	[1]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 20186;	
RX	MEDLINE=98007872; PubMed=9349716;	
RA	Paramontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,	
RA	van Loon A.P.G.M.;	
RT	"Cloning of the phytases from Emericella nidulans and the thermophilic fungus Talaromyces thermophilus.";	
RL	Biochim. Biophys. Acta 1353:217-223(1997).	
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE.	
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO- INOSITOL 1, 2, 4, 5, 6-PENTAKISPHOSPHATE + PHOSPHATE.	
CC	-1- MISCELLANEOUS: HAS POTENTIAL AS FOOD AND FEED ADDITIVE. IT CAN FACILITATE THE DEGRADATION OF PHYTN IN SOYBEAN AND OTHER SEEDS USED AS FOOD FOR MONOGASTRIC ANIMALS.	
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.	
DR	EMBL; U59802; AAB96873.1; ..	
DR	HSSP; P34752; 1IHP.	
DR	InterPro; IPR000560; ..	
DR	Pfam; PF00328; acid_phosphat; 1.	
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.	
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.	
KW	Hydrolase; Glycoprotein; Signal.	
FT	SIGNAL 1 14 POTENTIAL.	
FT	CHAIN 15 466 3-PHYTASE A.	
FT	ACT_SITE 79 79 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).	
FT	ACT_SITE 357 357 PROTON DONOR (BY SIMILARITY).	
FT	DISULFID 28 37 BY SIMILARITY.	
FT	DISULFID 68 410 BY SIMILARITY.	
FT	DISULFID 212 461 BY SIMILARITY.	
FT	DISULFID 261 278 BY SIMILARITY.	
FT		

[illegible]

Qy	235	TVARTSDATOLSPFCDLFADEW-QYDYLSL-KYIYGAGNPLCPAQGVCF-NELIARL	235
Db	268	TVSLTDDAHTLSPFCDLFADEWQYVNLISLDKYIYGGGNPLGPGVQGVWANELMARL	327
Qy	292	THSPVDODHTSTNHTLDSNPATEPLNATLYADFSDHNTWVSIFFAFLNGTKPLSTTSVE	351
Db	328	TRAPVHDHCVNTLDSNPATEPLNATLYADFSDHNSIVSIFWALGYNGLNGTAPLSQTSVE	387
Qy	352	SI-ETGYYAASVTVPFAARAYVEMMQCEAGGGGEGEKEPLRVRLVNDRVVPLRGGVDK	410
Db	388	SVSQTDGYAAATVPFAARAYVEMMQCRA-----EKEPLRVRLVNDRVVPLRGGVPTDK	440
Qy	411	LGRCCKLDDFEVGLSFARSGNWAECF 436	
Db	441	LGRCCKDAFVAGLSFAQAGGNWADCF 466	
RESULT	8		
O00107			
ID	O00107	PRELIMINARY;	PRT; 487 AA.
AC	O00107;		
DT	01-JUL-1997	(TrenBrel. 04, Created)	
DT	01-JUL-1997	(TrenBrel. 04, Last sequence update)	
DT	01-WAY-2000	(TrenBrel. 13, Last annotation update)	
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-		
DE	PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE		
DE	PHOSPHOHYDROLASE A).		
GN	PHYA.		
OS	Thielavia heterothallica.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
OC	Sordariates; Chaetomiaceae; Thielavia.		
OX	NCBI_TaxID=78579;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9717792; PubMed=9025298;		
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,		
RA	van Loon A.P.G.M.;		
RT	"The Phytase subfamily of histidine acid phosphatases: isolation of		
RT	genes for two novel phytases from the fungi Aspergillus terreus and		
RL	Myceliophthora thermophila".		
RL	Microbiology 143:245-252(1997).		
CC	-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE		
CC	FROM PHYTATE.		
CC	-!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MY-		
CC	INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.		
CC	-!- SUBCELLULAR LOCATION: SECRETED.		
CC	-!- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF		
CC	3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO		
CC	ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA		
CC	SHIFTED TO MORE ACIDIC PH VALUES.		
CC	-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.		
DR	EMBL; U59806; AAB52508.1; -		
DR	HSSP; P34752; 1IHP.		
DR	InterPro; IPR000360;		
DR	Pfam; PF00328; acid_phosphat; 1.		
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.		
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.		
KW	Hydrolase; Glycoprotein; Signal.		
FT	SIGNAL 1 ?	POTENTIAL.	
FT	CHAIN ?	487	3-PHYTASE A.
FT	DOMAIN ?	267	POLY-SER.
FT	DOMAIN 423	433	POLY-GLY.
FT	ACT_SITE 75	75	REQUIRED FOR BINDING SUBSTRATE (BY
FT			SIMILARITY).
FT	ACT_SITE 76	76	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE 368	368	PROTON DONOR (BY SIMILARITY).
FT	DISULFID 26	35	BY SIMILARITY.
FT	DISULFID 64	421	BY SIMILARITY.
FT	DISULFID 208	485	BY SIMILARITY.
FT	DISULFID 260	289	BY SIMILARITY.
FT	DISULFID 456	464	BY SIMILARITY.
FT	CARBOHYD 165	165	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 200	200	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 13.7%; Score 317.5; DB 3; Length 442;
 Best Local Similarity 25.8%; Pred. No. 8.7e-18;
 Matches 105; Conservative 67; Mismatches 158; Indels 77; Gaps 18;

QY 38 DVPKGRVTFVQVLSRHGARYPTSSKYSALIERIOK-NAT-FKGKYAFKTYNYTLG 95
 Db 55 DTPPHCEIEQAOLFMHRGEPFTKSGKQKFKYDKLKKANITDYKGLAFIEDLEYFVP 114
 QY 96 ADDLTPFGNQVNSGI---KF---YRRYKAL--ARNIVPEVRASGDRVIAAERFI 145
 Db 115 DSDNYELETTRGLYSGLLNAFKFYRLRYRSLVDTSSVLPFAAS-EDRVVDTFARSEF 173
 QY 146 EGFSAKLADPAHQASPVINVI---IPEGSYNNNTLDHGLC----TAFEDSTLGDREA 197
 Db 174 RGF-----FGPDYATSCSIQVNETDTSKANALTTKDN--CPTYNSSFYDYSFGDE--- 223
 QY 198 NETAFAPIRLEALPGVNLTDVNVLMDCPDFTVARTSDATQLSPPCDLFTADBEW 257
 Db 224 ---IFQREADRLNELSPGNITADITMGTYCAYETNVKGH-----SSFCDALESRAF 274
 QY 258 ---QYDYLSKLYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPP 314
 Db 275 IALQYNN-DVTKFYQFGPGYNMAGVAVYANATAKLQE----- 313
 QY 315 LNATLYADFSDHNTMVSIFPALLGYNGTKPLSTTSVESIETDGYAASVTPFAARAYVEM 374
 Db 314 -DGKLMFSFSDNDLLNYITALLGITDTE-LGTEDVDHFRS--FKTSELVPOGARLIIEK 369
 QY 375 MOCAGGGGGEKEPELVRVLVNDVRVPLHCGGVKGLCKLDDFVE 421
 Db 370 LNC-----SDTSFVRILNDKVPVPGCSSGPGYSCPLEDYLD 407

RESULT 10
 QY846 PRELIMINARY; PRT; 469 AA.
 ID QY846
 AC QY846;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE ACID PHOSPHATASE (EC 3.1.3.2).
 GN PHO3.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Kluyveromyces.
 OC Saccharomycetaceae; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2360/7;
 RA San Vicente A., Ferminan E., Dominguez A.;
 RT "Isolation and characterization of KlphO3 a gene encoding a
 constitutive acid phosphatase from Kluyveromyces lactis.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ007502; CAB46490.1;
 DR HSSP; P34755; IQEX.
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid.phosphat. 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase.
 FT CHAIN
 SQ SEQUENCE 469 AA; 52560 MW; 5C7ABF622CEA891C CRC64;

Query Match 12.9%; Score 298.5; DB 3; Length 469;
 Best Local Similarity 26.2%; Pred. No. 3.4e-16;
 Matches 117; Conservative 65; Mismatches 174; Indels 91; Gaps 21;

QY 21 GOYSPFFSLADESAISPDPKGRVTFVQVLSRHGARYPTSSKYSALIERIOKNAF 80
 Db 45 GO-GPHYDYPSQSGIPVEVPDQCTVEHVOMLARHGRIPTASKGLWIALWDLKE---F 100
 QY 81 KGKY-----AFLKTYNYTLGADDLTP-----FGENQMVNSGIKRYRYKALA 122

Query Match 55.0%; Score 1275.5; DB 3; Length 487;
 Best Local Similarity 56.3%; Pred. No. 5.2e-96;
 Matches 263; Conservative 52; Mismatches 119; Indels 33; Gaps 13;

QY 2 SHSCDTPD-GYQC-PEISHLMQYSPFSLADESAISPDPKGRVTFVQVLSRHGARYP 59
 Db 23 SRPCDTPDLGFCQGTATSHFGQYSPFVSE--LDASIPDDCEVTFAQVLSRHGARAP 80
 QY 60 TSSKSKYSALIERIOKNAFEGK-YAFKTYNYTLGADDLTPFGENQMVNSGIKFYRY 118
 Db 81 TLKRAASVVDLIRLHGAISYGPGEYELRYDYLGADELTRTGOQMVNSGIKFYRY 140
 QY 119 KALARNIVPVRASGDRVIAAERFIIEGFSAKLADPAHQASPI---NVIIEGSGYN 175
 Db 141 RALARKSIPFVRTAGODRVHSAENFTQGFHSALLADRGSTVRPTLPYDMVVIETAGAN 200
 QY 176 NTLHGLCTAFED---STLGGDAEANTAVFAPPIRLEA-LPGVNLTDVNVLMDCM 231
 Db 201 NTLHNDLCTAFEGPYSTIGDDAQYTLSTFAGPIARVANLPGANLTDADTVALMDLC 260
 QY 232 PPDVARTS-----DATQLSPPCDLFTADBEW-QYDYLSKLYGYGAGNPLGP 278
 Db 261 PFTVASSSDPATADAGGNGRPLSPFCRLFSSEWRAYDYLSVQKMGYGPNGPLGP 320
 QY 279 AQGVGF-NELIARLTHSPVQDHTSTNHTLDSNPATFLNATLYADFSDHNTMVSIFFAIG 337
 Db 321 TQGVGEVNEELLARLGVPRDGTSTNRTLDGDPRTFPLGRPLIYADFSDHNDMMGLGALG 380
 QY 338 LYGTYPLSTTS-VESETDGYAASVTPFAARAYVEMMOCAGGGGGBG-----EKEP 390
 Db 381 AYDGVPLDKTARRDPEELGYAASVAVPFAARIYVEKMRCSGGGGGGGEGREGKDEE 440
 QY 391 LVRVLVNDVRVPLHCGGVKGLCKLDDFVEGLSFARSGNNAECFA 437
 Db 441 MVRVLVNDVRVMTLKGCCADRGCTLERFTIESMAFARGNGKWDLCFA 487

RESULT 9
 ID 074677 PRELIMINARY; PRT; 442 AA.
 AC 074677;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE (EC 3.1.3.2).
 GN PHO1.
 OS Pichia angusta (Yeast) (Hansenula polymorpha).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC34436;
 RC MEDLINE=98386672; PubMed=9720203;
 RA "Cloning and characterization of the gene encoding a repressible acid
 phosphatase (PHO1) from the methylotrophic yeast Hansenula
 polymorpha.";
 RT Appl. Microbiol. Biotechnol. 50:77-84(1998).
 RL EMBL; AF051161; AAC2537.1;
 DR HSSP; P34755; IQEX.
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid.phosphat. 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 FT CHAIN
 SQ SEQUENCE 442 AA; 49370 MW; 7087D91A85B053C1 CRC64;

```

Db 101 OCQYNGPMVEVNDYEFFVSNKYEDQLTNSDVPSPYAGAKTAQHLGKYIAYNYGDLF 160
QY 123 RNIVPFVRASGDRVIAAEKFIQFQSAKLADPAHQASPVINYV---IIPEG---SGYNNT 177
Db 161 SDSNP-VFTSSGRV-----HOTAKYVVSLSSEELDQLDQLIQENETSGANSL 209
QY 178 LDHGLCTAFEDSTLGDAAEANTAVFAPPRIAR-LEALPGVNLT--DEDVNVNMDMCPED 234
Db 210 TPADSCMTY-NGDLGDEYFENATLPYLTIDKNRMWKNKSNLNTLEHDDLELLVDWCAFE 268
QY 235 TVARTSATQSLSPDCDLFTADEW-QYDYLSL-KYGYGAGNPL-GPAQGVGFNELIARL 291
Db 269 TNVRGSSAV-----CDLFRNDLVAYSYYANVNNYRRGAGNPMSPGIVLVNYSYNLL 323
QY 292 THSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTWVSIFALGLY-NGTKPLSTTSV 350
Db 324 TOADELND-----KVLFSFSDHTDIOQFISALGLDNGVTEYSILDQV 365
QY 351 --ESIEFDGYAASWTVPFAARAYVEMMOCBAGGGGEGEKEPLRVLVNDRVVPPLHGGCV 408
Db 366 DFQNIQ---OLSWVTMPMGRIETFKLC-----GNASVRYIINDVIIPVPGCTS 412
QY 409 DKLGRCKLDDF-----VEGLSFARS 428
Db 413 GPGFSCPIEDFDYITNRLNGIDYVSS 439

RESULT 11
O60172
ID O60172 PRELIMINARY; PRT; 463 AA.
AC O60172;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE PROBABLE ACID PHOSPHATASE (EC 3.1.3.2).
GN SPBC2187.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX STRAIN=972;
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; AL023286; CAA18863.1; ..
DR HSSP; P34755; 10FX.
DR InterPro; IPR000560; ..
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein.
FT ACT SITE 69
FT ACT SITE 340
FT ACT SITE 340
FT CARBOHYD 98
FT CARBOHYD 98
FT CARBOHYD 104
FT CARBOHYD 104
FT CARBOHYD 221
FT CARBOHYD 221
FT CARBOHYD 324
FT CARBOHYD 324
FT CARBOHYD 439
FT CARBOHYD 439
FT CARBOHYD 458
FT CARBOHYD 458
SQ SEQUENCE 463 AA; 52758 MW; 6C41AF422CD624A CRC64;

```

```

Query Match 12.3%; Score 285; DB 3; Length 463;
Best Local Similarity 24.9%; Pred. No. 4.2e-15;
Matches 113; Conservative
QY 23 YSPFSLADESAISPDVPGKGRVTFQVLSRHGARYPTS--SKSKKYSA-----L 70

```

```

Db 43 HEFYFDGLDSSA-----FPTCEIQVHLLRHGSRNPTGDTVATDVIYSOYLNNEQEKLL 97
QY 71 IERIQKNATF-GKVAFLKTYNYTL---GADDLTPFGENQMVNSGIKFYRRYKALARNIV 126
Db 98 NSGIPVNFYSYPENPLCFIKQWTPVIDAENADQLSSRGLRLEFDLGRQLYQRYKLFDSV 157
QY 127 PFVRASGSDRVIAAEKFIQFQSAKLADPAHQASPVINYV---IIPEG---SGYNNTLDHGLCT 184
Db 158 YDINTAEQERVVESAKWFIYGLFGDKMYEKN-----FILISKEGAAGANSLSYNACP 211
QY 185 AFEDSTL-----GDDAEANTAVFAPPRIARLEAL--PGVNLTDDEVDVNMDCPDPDVAR 238
Db 212 VFKNFNHKNKNTDAHAHVNRNIFIEPIVNRLLAKYFDSSYKLTINDVRSIFICEYIAIK 271
QY 239 TSDATQLSPDCDLFTADEW-QYDYLSL-KYGYGAGNPL-GPAQGVGFNELIARLTHSPVQ 297
Db 272 DH-----SDFCSIFTPSEFLNFEYDSLDL-QAYGGG-----PVS 304
QY 298 DHTST-----NHTLDS--NPATFPLNATLYADFSDHNTWVSIFALGLY---NGTKPLS 346
Db 305 EWASTLGGAYINNLADSLRNVTNPDQKRVFLAETHDSNIIPVEAALGFFPDITPQPLP 364
QY 347 TTSVESIETDGYAASWTVPFAARAYVEMMOCBAGGGGEGEKEPLRVLVNDRVVPPLHGC 406
Db 365 TD--KNIITYSOKTSSSFVPFAGNLITELFFC-----SDSKYYVRHLVNOQVYPLDLC 414
QY 407 GVDKILGR-----CKLDDFVEGLSFARSNGNNAECF 436
Db 415 GYGPSGTSGLCELOAYLNSPIRANSTNSGISIF 448

RESULT 12
Q9VVW72
ID Q9VVW72 PRELIMINARY; PRT; 460 AA.
AC Q9VVW72;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE MIPPI PROTEIN.
GN MIPPI OR CG4123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey L.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

```

[illegible]

Query Match	8.3%;	Score 192.5;	DB 5;	Length 453;
Best Local Similarity	23.4%;	Pred. No. 1.5e-07;		
Matches 101;	Conservative 64;	Mismatches 174;	Indels 93;	Gaps
QY 24 SPFFSLA--DESAISPDVPGKGRVTFVOVLSRHGARYPTSSKKKYSALIERIOKNAFK 81	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 43 TYRAIANYDE---TPPKYACGHPTRIWTITRHGTRNPSESILVLAQNRLSEIKRILDQ 99	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 82 GRYAF-----LKYNYT-LGADD---LTPFGENQMVNSGIKFRYRYKALARNIVP-- 127	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 100 TRPPICTAELEKLRQWHHNLNATEDEKLLVAEGEDELIELAERMRQRRFPDLLPELYNPE 159	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 128 --FVRASGSDRVIASAEKIEG-----FOSAKLADPAHQASPVINVIIEGSGYNNLTLD 179	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 160 WYFYKTYATQRTLSAESFATGLFGRHRIHTRVYPPLHE-DPVLRY--KGGCKWKT-- 214	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 180 HGLCTAFEDSTLGDGAENAFVAFV-PIPARLE-----ALPGVNTLDEGVNLMMDMCP 232	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 215 -----DVKNPETLVNARFLAEPOMOSAEQVRSSTRLP--DLQPEDVQLMTVCA 264	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 233 FDTVART-----SDATQSLSPFCDLFTADEWQ-YDYLSQSLKY--GYGAGNELPLGAQGV 282	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 265 FETAHRRPHRDSGSKSSVSWCNCFFDVAAEALFEFDELEYIWNQGY----- 313	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 283 GFNELIARLTHSPVDH-----TSTNHITLDSNPATFPPLNATLYADFSDHNTWVSIFPFG 338	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 314 --YELTHRIACPAIDMFAISSEETQRRR-----ANATLY--FTHSGTLLKLAHLGL 364	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 339 YNGTKPLSTTSVESIETDGVAASTVTPFAARAYVEMMQCEAGGGGEGEKEPLRVLVND 398	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 365 ARDNKPLTHKHFAERL--NRTSOIDAFATNLALFLRYDCKG-----NPQVLVHOE 414	!! : : : !	!! : : : !	!! : : : !	!! : : : !
QY 399 RVPLHGGGVDK 410	!! : : : !	!! : : : !	!! : : : !	!! : : : !
Db 415 RWRLPGCPQDK 426	!! : : : !	!! : : : !	!! : : : !	!! : : : !
RESULT 15				
Q9W438				
ID Q9W438	PRELIMINARY;	PRT;	453 AA.	
AC Q9W438;				
DT 01-MAY-2000 (Tremblrel. 13, Created)				
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)				
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)				
DE M1P2 PROTEIN.				
GN M1P2 OR CG4317.				
OS Drosophila melanogaster (Fruit fly).				
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC Ephydroidea; Drosophilidae; Drosophila.				
NCBI_TaxID=7227;				
RN [1]				
RC SEQUENCE FROM N.A.				
RP STRAIN=BERKELEY;				
RX MEDLINE=20196006; PubMed=10731132;				
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA Ananietos P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,				
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,				
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,				
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				

Search completed: October 26, 2001, 15:17:25
Job time: 3976 sec

us-09-488-265-27.rspt

Sat Oct 27 15:25:44 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:03 ; Search time 57.24 Seconds
(without alignments)
462.835 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDTVDGQCPEISHLW.....DFVEGLSFARSGGNWAECA 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*

- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2321	100.0	437	21	AA20531
2	2321	100.0	437	21	AA20531
3	2115	91.1	441	21	AA20523
4	2115	91.1	467	21	AA20524
5	2115	91.1	467	21	AA20526
6	2099	90.4	467	21	AA20533
7	2093	90.2	467	20	AA20533
8	2093	90.2	467	21	AA20527
9	2093	90.2	467	21	AA20534
10	2093	90.2	467	21	AA20534
11	2067	89.1	467	21	AA20531

12	2064.5	88.9	424	21	AA20536
13	2064.5	88.9	424	21	AA20536
14	2061	88.8	467	21	AA20532
15	2044	88.1	467	21	AA20532
16	2044	88.1	467	21	AA20526
17	2044	88.1	467	21	AA20526
18	2044	88.1	467	21	AA20526
19	2020	87.0	441	21	AA20514
20	2020	87.0	467	20	AA20514
21	2020	87.0	467	20	AA20515
22	2020	87.0	467	21	AA20515
23	2014	86.8	467	21	AA20515
24	2013	86.7	467	20	AA20515
25	2013	86.7	467	20	AA20515
26	2005	86.4	467	20	AA20515
27	2004	86.3	467	20	AA20515
28	1997	86.0	431	21	AA20515
29	1997	86.0	431	21	AA20515
30	1954	84.2	467	21	AA20530
31	1954	84.2	467	21	AA20530
32	1922	82.8	467	20	AA20530
33	1922	82.8	467	20	AA20530
34	1919	82.7	467	21	AA20530
35	1797.5	77.4	467	21	AA20529
36	1777.5	76.6	467	20	AA20529
37	1777.5	76.6	467	20	AA20529
38	1777.5	76.6	467	21	AA20528
39	1762.5	75.9	440	21	AA20528
40	1762.5	75.9	440	21	AA20528
41	1762.5	75.9	465	19	AA20528
42	1759.5	75.8	440	21	AA20528
43	1759.5	75.8	440	21	AA20528
44	1754.5	75.6	440	21	AA20528
45	1754.5	75.6	440	21	AA20528

ALIGNMENTS

RESULT 1	
ID	AA20525 standard; Protein; 437 AA.
XX	
AC	AA20525;
XX	
DT	05-DEC-2000 (first entry)
XX	
DE	Consensus phytase 11 SEQ ID NO:27.
XX	
KW	Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW	temperature stability; pH profile; temperature profile; reaction rate;
KW	specific activity; substrate specificity; substrate cleavage pattern;
KW	substrate binding; position specificity; phytate degradation rate;
XX	food; feed; phytate; manure.
OS	Synthetic.
XX	
PN	WO2000/43503-A1
PD	27-JUL-2000.
XX	
PF	21-JAN-2000; 2000WO-DK00025.
XX	
PR	22-SEP-1999; 99DK-0000092.
PR	21-SEP-1999; 99DK-0001340.
XX	
PA	(NOVO) NOVO NORDISK AS.
XX	
PI	Lehmann M;
XX	
DR	WPI; 2000-491161/43.
XX	
PT	Novel phytases with improved properties such as temperature stability,

Db 1 nshscdtvdygcpeishlwgqyspfsladesaispdkgrvtfvqlsrhgarypt 60
 QY 61 SSKSKYSALIERIOKNATFKGKYAFKTYNTLIGADDLTPFGENQMVNSGKIFRYRYKA 120
 Db 61 sskskysalieriknafcgykafktyntlygaddltfgenqmvnsqikfryryka 120
 QY 121 LARNIVPFVRASGDRVIAAEKFTGFSQAKLADPAHQASPVINVIIEGSGYNNTLDH 180
 Db 121 larnivpfvrasgdrvriaaeekfiegfsakladpahqaspviniiepgsgynntldh 180
 QY 181 GLCTAFEDSTGLDDEAFNFTAVFAPPFARLEALPGVNLTDDEVDVNLMDMCPDFTVARTS 240
 Db 181 glctafedstlgddaeafnftavfappfarrlealpgvnltdedvnlmdmcpdftvarts 240
 QY 241 DATQLSPFCDLTADWQYDYLQSLKYYGAGNPLGPAQGVGFNELLARLTHSPVQDHT 300
 Db 241 datqlspfcldtadewdydylqslkyygagmplgpaqgvgnellarlthspvqdh 300
 QY 301 STNHTLDSNPATFPLNATLYADFSHDNTMVSIFFALGLYNGTKPLSTTSVESIETDGYAA 360
 Db 301 stnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsvesietdgyaa 360
 QY 361 SWTPFPAARAYVEMMOCEAGGGGEGEKEPLVRVLRVNDRVVPLHGGCVDKLGRCKLDDFV 420
 Db 361 swtpfpaarayvemmoceaggggggekeplvrvlrvlndrvvplhggcvdklgrckldfv 420
 QY 421 EGLSFARSGGNWAECEFA 437
 Db 421 eglsgfarsgggnwaecefa 437

RESULT 3
 AAB20523
 ID AAB20523 standard; Protein: 441 AA.
 AC AAB20523;
 DT 05-DEC-2000 (first entry)
 XX Consensus phytase 10 (Fcp10) SEQ ID NO:24.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.
 OS
 XX WO200043503-A1
 XX 27-JUL-2000.
 XX
 XX 21-JAN-2000; 2000WO-DK00025.
 XX 22-JAN-1999; 99DK-0000092.
 XX 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 XX Lehmann M;
 XX WPI; 2000-491161/43.
 XX
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX
 PS Example 2; Fig 4a-d; 240pp; English.
 XX
 XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.

CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX
 XX Sequence 441 AA;

Query Match 91.1%; Score 2115; DB 21; Length 441;
 Best Local Similarity 93.5%; Pred No. 3.1e-198;
 Matches 419; Conservative 2; Mismatches 9; Indels 18; Gaps 10;
 QY 1 NSHSCDTVD-GYOC-PEISHLMGQYSPFSLADESAISPDKGRVTFVQVLSRHGARY 58
 Db 1 nshscdtvdygcpeishlwgqyspfsladesaispdkgrvtfvqlsrhgary 60
 QY 59 PTSSSKYSALIERIOKNAT-FKGKYAFKTYNTLIGADDLTPFGENQMVNSGKIFRYR 117
 Db 61 ptssskysalieriknatafkgyafktyntlygaddltfgenqmvnsqikfryr 120
 QY 118 YKALARNIVPFVRASGDRVIAAEKFTGFSQAKLADPA---HQASPVINVIIEGSGY 174
 Db 121 ykalarivpfvrasgdrvriaaeekfiegfsakladpahqaspviniiepgsgy 180
 QY 175 NNTLDHGLCTAFEDSTGLDDEAFNFTAVFAPPFARLEA-LPGVNLTDDEVDVNLMDMCP 233
 Db 181 nntldhglctafeseldgdeafnftavfappfarrleahlpvnltdedvnlmdmcpf 240
 QY 234 DTVARTSDATQLSPFCDLTADWQYDYLQSLKYYGAGNPLGPAQGVGFNELLAR 290
 Db 241 dtvartsdatlspfcldtadewdydylqslkyygagmplgpaqgvgnellar 300
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFALGLYNGTKPLSTTSV 350
 Db 301 lthspvqdhstnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsv 360
 QY 351 ESI-ETDCGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRVLRVNDRVVPLHGGCV 409
 Db 361 esieetdgyaaswtvpfaarayvemmoceaggggggekeplvrvlrvlndrvvplhggcv 413
 QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 414 klgrckrddfveglsgfarsgggnwaecefa 441

RESULT 4
 AAB20524
 ID AAB20524 standard; Protein: 467 AA.
 AC AAB20524;
 DT 05-DEC-2000 (first entry)
 XX Consensus phytase 10 SEQ ID NO:26.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.
 OS
 XX WO200043503-A1.
 XX 27-JUL-2000.

PF	21-JAN-2000; 2000WO-DK00025.	AA69566	ID	AA69566 standard; protein; 467 AA.
XX		XX	AC	AA69566;
PR	22-JAN-1999; 99DK-0000092.	XX	DT	19-APR-2000 (first entry)
PR	21-SEP-1999; 99DK-0001340.	XX	DE	Phytase-10, a consensus phytase.
XX		XX	KW	Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
PA	(NOVO) NOVO NORDISK AS.	KW	KW	thermostable; animal feed; monogastric animal; phytate phosphorus;
XX		XX	XX	phosphate availability; consensus; phytase-10.
PI	Lehmann M;	OS	XX	Aspergillus terreus 9A1.
XX		OS	OS	Aspergillus terreus cbs16.46.
XX		OS	OS	Aspergillus niger var. awamori.
DR	WPI; 2000-491161/43.	OS	OS	Aspergillus fumigatus str. NRRL3135.
DR	N-PSDB; AAA73232.	OS	OS	Aspergillus fumigatus ATCC13073.
XX		OS	OS	Aspergillus fumigatus ATCC32722.
PT	Novel phytases with improved properties such as temperature stability,	OS	OS	Aspergillus fumigatus ATCC58128.
PT	pH stability and substrate specificity, for use in pharmaceuticals and	OS	OS	Aspergillus fumigatus ATCC26906.
PT	compound foods and feeds	OS	OS	Aspergillus fumigatus ATCC32239.
XX		OS	OS	Emmericella nidulans.
PS	Claim 1; Fig 5a-c; 240pp; English.	OS	OS	Talaromyces thermophilus ATCC20186.
XX		OS	OS	Myceliophthora thermophila.
XX		OS	OS	Paxillus involutus NN005693.
CC	The present invention describes improved phytases, preferably with	OS	OS	Trametes pubescens NN9343.
CC	increased thermostability, and methods for producing them. The methods	OS	OS	Agaricus pediades NN009289.
CC	can be used for producing phytases with improved properties e.g.	OS	OS	Peniophora lycii NN006113.
CC	temperature stability, pH stability, pH profile, temperature profile,	OS	OS	Thermomyces lanuginosa.
CC	specific activity, substrate specificity, substrate cleavage pattern,	OS	OS	Synthetic.
CC	substrate binding, position specificity, the velocity and level of	XX	PH	Key
CC	release of phosphate from corn, reaction rate, phytase degradation rate,	XX	FT	Peptide
CC	and end level of released phosphate. The phytases can be used to produce	XX	FT	1..26
CC	pharmaceutical compositions or compound food or feeds. The feed can be	XX	FT	/note= "Phytase signal peptide from Aspergillus terreus
CC	used to reduce levels of phytate in animal manure, by converting it	XX	FT	/note= "Mature phytase-10 consensus"
CC	into lower inositol phosphates and/or inositol and inorganic phosphate.	XX	XX	Protein
CC	The present sequence represents a phytase sequence from the present	XX	XX	27..467
CC	invention.	XX	XX	/note= "Mature phytase-10 consensus"
XX		XX	XX	EP969089-A1.
SQ	Sequence 467 AA;	XX	XX	05-JAN-2000.
		XX	XX	23-JUN-1999; 99EP-0111949.
		XX	XX	29-JUN-1998; 98EP-0111960.
		XX	XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
		XX	XX	Brugger R, Lehmann M, Wyss M;
		XX	XX	WPI; 2000-099429/09.
		XX	XX	N-PSDB; AAZ59642.
		XX	XX	New stabilized enzyme formulation, useful for feed compositions for
		XX	XX	monogastric animals
		XX	XX	Example 4; Fig 17; 101pp; English.
		XX	XX	The invention relates to a novel stabilised dry or liquid enzyme
		XX	XX	formulation, comprising phytase (myo-inositol hexakisphosphate
		XX	XX	phosphohydrolase) and one or more stabilising agents including
		XX	XX	xylytol or ribitol; polyethylene glycols with a molecular weight of 600
		XX	XX	to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
		XX	XX	glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
		XX	XX	The stabilised phytase formulation is used in a method for preparing a
		XX	XX	feed composition for monogastric animals (e.g., pigs, poultry) and
		XX	XX	provides a monogastric animal with its dietary requirements of
		XX	XX	phosphorus. Although a large amount of phosphate is present in animal
		XX	XX	feed in the form of phytate phosphorus, monogastric animals are unable
		XX	XX	to utilise this form of phosphate, resulting in the addition of extra
		XX	XX	phosphate to the feed of such animals. Phytase enhances the nutritional
		XX	XX	value of plant material without the need for adding additional phosphate

Query Match	91.1%;	Score	2115;	DB	21;	Length	467;
Best Local Similarity	93.5%;	Pred.	No. 3.4e-198;				
Matches	419;	Conservative	2;	Mismatches	9;	Indels	18;
						Gaps	10;
QY	1	NSHSCDVTVD-GYOC-PEISHLQGVSPFFSLADESAISPDVPGKGRVTFVQVLSRHGARY	58				
Db	27	nshscdvtvggycfeishlwgqyspffsladesaispdpkgkgrvtvqvlshrghary	86				
QY	59	PTSSKSKYSALIEROKNAT-PKGKAYFLKTYNTLGADLTTPFCENOMVNSGKIFVRR	117				
Db	87	ptsskskysallieaigknatafkgyarfktyntlgadlttppfcqgmvasgkifvrr	146				
QY	118	YKALARNVFPVRASGDRVIAAEKFIQFQSAKLADPA---HQASPVINVTIIPGSGY	174				
Db	147	ykalarvfpvragsdrviasaekfiegfsakladpaganphqaspvlnvlipegagy	206				
QY	175	NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPTARLEA-LPGVNLTDENVNLMDCPF	233				
Db	207	ntntldhglctafeseelgddveanftavfappirarleahlpqvnltdeenvnldmcpf	266				
QY	234	DTVARTSDATQLSPFCDLTADWE-QVDYLOSL-KYGYGAGNPLGPAQGVF-NELIAR	290				
Db	267	dtvartsdatqlspfcldtfdewiqdyqlsgkygygagnpplpagaqgvfneliar	326				
QY	291	LTHSPVDQHTSTNHTLSDNPATFPLNATLYADFSDHNTWVSIFALGLYNGTKPLSTSV	350				
Db	327	lthspvqdhstnhtlgsnpatfplnatlyadfsdntmvsifalglyngtkplstsv	386				
QY	351	ESI-ETDGYAASVTWPAARAYVEMMOCEAGGGGEGEKEPLVRVLVNDRVPLHGGVD	409				
Db	387	esieetdgyaaswtvtaarayvemmqcea-----ekeplvrvlvndrvplhggvd	439				
QY	410	KLGRCKLDDFVEGLSFARSNGNWAECFA	437				
Db	440	klgrckrddffveglsfarsngnweecfa	467				

CC to the feed. The level of phosphate pollution in the environment is
CC reduced by adding phytase to animal feed, as the animal can make use of
CC the inorganic phosphate liberated from phytate phosphorus using the
CC enzyme. The phytase formulation of the invention has an improved
CC thermostability and can therefore remain stable during long-term storage
CC and can withstand feed processing methods such as extrusion, expansion
CC and pelleting. The present sequence represents a consensus sequence,
CC designated phytase-10, which was derived from the mature phytase
CC sequences from a variety of fungi (AAV69544-V69546, AAV69548-V69556,
CC AAV69564) and the Basidiomycetes phytase consensus AAV69563 and
CC additionally contains the Aspergillus terreus cbs116.46 signal peptide at
CC the N-terminus.
XX
SQ Sequence 467 AA;

Query Match 91.1%; Score 2115; DB 21; Length 467;
Best Local Similarity 93.5%; Pred. No. 3.4e-198;
Matches 419; Conservative 2; Mismatches 9; Indels 18; Gaps 10;
QY 1 NSHSCDWD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
DB 27 nshscdtdvgyqcfpeishlwgyqspffsladesaispdvpgkcrvtfvqvlshrghary 86
QY 59 PTSSKSKKYSALIERIOKNAT-FKGKYAFKTYNTLGADDLTPFGENQMVNSGKIFR 117
DB 87 ptsskskksalierioknat-fkgkyafktyntlgadddltpfgeqgmvnsgikfyr 146
QY 118 YKALARNIVPFVRASGSDRVIASAEKFIQFQSAKLADPA---HOASPVNVIIPGSGY 174
DB 147 ykalarkivpfvrasgsdrviasaeekfiqfqsakladpganphgaspvniilpegag 206
QY 175 NNTLDHGLCTAFEDSLGDDAEANFTAVFAPPIRARLEA-LPGVNLTDDEVDVNLDMCP 233
DB 207 nntldhglctafeeslgddveanftavfappirarleahlpvnltdedvnlmdmcpf 266
QY 234 DTVARTSDATQLSPFCDFLTADSW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 267 dtvartsdatqlspfcdfthdewigdyqlsgkygygagnglpagdgvgfvnelliar 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFAFGLYNGKPLSTTSV 350
DB 327 lthspvgdhtstnhtldsnpatfplnatlyadfsdntmvsiffalgyngkplsttsv 386
QY 351 EST-ETDGYAASWTVPFAARAVVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGCQVD 409
DB 387 esieetdgyaaswtvpfaarayvemmqcea-----ekeplrvlvndrvvplhgcgvd 439
QY 410 KLGRCCKLDDFVEGLSFARSGGNWAECEFA 437
DB 440 klgrckrddfveglfsarsggnweecfa 467

RESULT 6
AAB20533
ID AAB20533 standard; Protein: 467 AA.
XX
AC AAB20533;
XX
DT 05-DEC-2000 (first entry)
XX
DE Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95.
XX
KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure.
XX
OS Synthetic.
XX
PN W0200043503-A1.
XX

PD 27-JUL-2000.

XX
PF 21-JAN-2000; 2000WO-DK00025.
XX
PR 22-JAN-1999; 99DK-0000092.
PR 21-SEP-1999; 99DK-0001340.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Lehmann M;
XX
WPI: 2000-491161/43.
N-PSDB; AAA73292.

Novel phytases with improved properties such as temperature stability,
pH stability and substrate specificity, for use in pharmaceuticals and
compound foods and feeds -
Disclosure: Fig 24a-c; 240pp; English.

The present invention describes improved phytases, preferably with
increased thermostability, and methods for producing them. The methods
can be used for producing phytases with improved properties e.g.
temperature stability, pH stability, pH profile, temperature profile,
specific activity, substrate specificity, the velocity and level of
substrate binding, position specificity, the velocity and level of
release of phosphate from corn, reaction rate, phytate degradation rate,
and end level of released phosphate. The phytases can be used to produce
pharmaceutical compositions or compound food or feeds. The feed can be
used to reduce levels of phytate in animal manure, by converting it
into lower inositol phosphates and/or inositol and inorganic phosphate.
The present sequence represents a phytase sequence from the present
invention.

SQ Sequence 467 AA;

Query Match 90.4%; Score 2099; DB 21; Length 467;
Best Local Similarity 92.6%; Pred. No. 1.2e-196;
Matches 415; Conservative 4; Mismatches 11; Indels 18; Gaps 10;
QY 1 NSHSCDWD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
DB 27 nshscdtdvgyqcfpeishlwgyqspffsladesaispdvpgkcrvtfvqvlshrghary 86
QY 59 PTSSKSKKYSALIERIOKNAT-FKGKYAFKTYNTLGADDLTPFGENQMVNSGKIFR 117
DB 87 ptsskskksalierioknat-fkgkyafktyntlgadddltpfgeqgmvnsgikfyr 146
QY 118 YKALARNIVPFVRASGSDRVIASAEKFIQFQSAKLADPA---HOASPVNVIIPGSGY 174
DB 147 ykalarkivpfvrasgsdrviasaeekfiqfqsakladpganphgaspvniilpegag 206
QY 175 NNTLDHGLCTAFEDSLGDDAEANFTAVFAPPIRARLEA-LPGVNLTDDEVDVNLDMCP 233
DB 207 nntldhglctafeeslgddveanftavfappirarleahlpvnltdedvnlmdmcpf 266
QY 234 DTVARTSDATQLSPFCDFLTADSW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 267 dtvartsdatqlspfcdfthdewigdyqlsgkygygagnglpagdgvgfvnelliar 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFAFGLYNGKPLSTTSV 350
DB 327 lthspvgdhtstnhtldsnpatfplnatlyadfsdntmvsiffalgyngkplsttsv 386
QY 351 EST-ETDGYAASWTVPFAARAVVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGCQVD 409
DB 387 esieetdgyaaswtvpfaarayvemmqcea-----ekeplrvlvndrvvplhgcgvd 439
QY 410 KLGRCCKLDDFVEGLSFARSGGNWAECEFA 437
DB 440 klgrckrddfveglfsarsggnweecfa 467

Db 267 dtvartsatqlspfcldlthdewiqdyqlslogkygygagpnlpagqvgfvneliar 326
 QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSHDNTMWSIFALGILYNGTKPLSTTSV 350
 Db 327 lthspvqdhstnhtldsnpatfnatlyadfshtntmwsifalglyngtkplstsv 386
 QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRVLVNDVRVPLHGCGVD 409
 Db 387 esieetdgysaswtvpfaarayvemmqcea-----ekeplvrvlvndrvrplhgcvd 439
 QY 410 KLCRCKLDDFVEGLSFARSGGNWAECA 437
 Db 440 kigrckrddfvglsgfarsggnweecfa 467

RESULT 8

AAB20527
 ID AAB20527 standard; Protein: 467 AA.

AC AAB20527;

XX 05-DEC-2000 (first entry)

Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.

Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 temperature stability; pH profile; temperature profile; reaction rate;
 specific activity; substrate specificity; substrate cleavage pattern;
 substrate binding; position specificity; phytate degradation rate;
 food; feed; phytate; manure.

OS Synthetic.

XX W0200043503-A1.

PN WC200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI: 2000-491161/43.

DR N-PSDB; AA73234.

Novel phytases with improved properties such as temperature stability,
 pH stability and substrate specificity, for use in pharmaceuticals and
 compound foods and feeds -

Claim 4; Fig 8a-c; 240pp; English.

The present invention describes improved phytases, preferably with
 increased thermostability, and methods for producing them. The methods
 can be used for producing phytases with improved properties e.g.
 temperature stability, pH stability, pH profile, temperature profile,
 specific activity, substrate specificity, substrate cleavage pattern,
 substrate binding, position specificity, the velocity and level of
 release of phosphate from corn, reaction rate, phytate degradation rate,
 and end level of released phosphate. The phytases can be used to produce
 pharmaceutical compositions or compound food or feeds. The feed can be
 used to reduce levels of phytate in animal manure, by converting it
 into lower inositol phosphates and/or inositol and inorganic phosphate.
 The present sequence represents a phytase sequence from the present
 invention.

XX Sequence 467 AA;

SQ

RESULT 7
 AAY43170
 ID AAY43170 standard; Protein: 467 AA.
 XX
 AC AAY43170;
 XX
 DT 06-JAN-2000 (first entry)
 DE Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.
 XX phytase; animal feed preparation; thermostable phytase; transgenic plant;
 KW consensus sequence.
 KW Synthetic.
 OS
 XX W09948380-A1.
 PN
 PD 30-SEP-1999.
 XX
 PF 22-MAR-1999; 99WO-DK00154.
 XX
 PR 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 PR 22-JAN-1999; 99DK-0000093.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Petersen S;
 PI
 XX WPI: 1999-591030/50.
 DR N-PSDB; AA231521.
 DR
 XX Preparing animal feed using a thermostable phytase

Example 3; Fig 10; 71pp; English.

This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
 The invention relates to a process for preparing animal feed by
 agglomerating feed ingredients with a thermostable phytase, which is
 added before or during agglomeration. The thermostable phytase is useful
 for expression in transgenic plants. These plants are useful in the
 preparation of animal feed itself. The thermostable phytase allows animal
 feed to be produced more efficiently. In addition to improved
 phytase-expressing transgenic plants, these plants provide a feed
 ingredient and a feed additive (phytase) simultaneously.

Sequence 467 AA;

Query Match 90.2%; Score 2093; DB 20; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;
 QY 1 NSHSCDFVD-GYOC-PEISHLWGOYSPFFSLADESAISDPVPGKCRVTFVQVLSRHGARY 58
 Db 27 nshscdctvdggyqcfpeishlwgtyspffsladesaisdpvpgkcrvtfvqvlstrhgary 86
 QY 59 PTSSSKRKYKYSALIERIOKNAT-FPGKYAFUKTYNYTLGADDLTPFGENQMVNSGKIFYR 117
 Db 87 ptssaskaysallcaiqknatafkgyafiktyntlgadddltpfgeqmvnsgkifyr 146
 QY 118 YKALARNIVFVRASGDRVTSASEKIEFGQSAKLADPA---HQASPVINVIPIEGSGY 174
 Db 147 ykalarivpfiragsdrviassekiefqskladpaganphqaspvinviipegagy 206
 QY 175 NTLDHGLCTAFEDSTLGDAAENFTAVFAPPTRARLEA-LPGVNLTDDEVNLMDCPF 233
 Db 207 ntlidhglctafeseldgdeanftavfappirarleahlpvgnltdevnldmldmcpf 266
 QY 234 DTVARTSDATOLSPFCDLFTADEW-QYDYLOSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290

Query Match 90.2%; Score 2093; DB 21; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLMGQYSPFFSLADESAISPDPVKGCRVTFVOVLSRHGARY 58
 DB 27 nshscdvtvgyqcfpeishlwgtyspffsladesaispdpvkgrvtfvqvlshrghary 86
 QY 59 PTSSSKKYSALIERIOKNAT-FKGKYAFKTYNTLGADDLTPFGENOMVNSGKIFYRR 117
 DB 87 ptssaskaysalieraiknatafkgkyafktyntlgadddtpfgeqmqvnsqikfyrr 146
 QY 118 YKALARNIVPFVRASGSDRVIASAEKFTIEGFSQAKLADPA---HQASPVINVIIEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaeekfiegfsqakladpaganphqaspinviiepegagy 206
 QY 175 NNTLDHGLCTAFEDSTLGDDEANFTAVFAPPPIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafeeseldgdeanftavfappirarleahpgvnltdedvnlmdmcpf 266
 QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatsqspfcldlftdewiqydyqlsgkygygagnpbgagvgvfnelliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDMTMSIFFALGLYNGTKPLSTTSV 350
 DB 327 lthspvqdhstnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsv 386
 QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEGEKEPLVRLVNDVRVPLHGGCVD 409
 DB 387 esieetdgyaswtvpfaarayvemmqcea-----ekeplvrvlndrvrvplhgcgvd 439
 QY 410 KLGRCKLDDFVEGLSFARSGGNNAECFA 437
 DB 440 klgrckrddfveglstarsgnweecfa 467

RESULT 9
 AAB20534
 ID AAB20534 standard; Protein; 467 AA.
 AC AAB20534;
 XX 05-DEC-2000 (first entry)
 DT Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.
 OS
 PN WO200043503-A1.
 XX 27-JUL-2000.
 XX 21-JAN-2000; 2000WO-DK00025.
 XX 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PA Lehmann M;
 XX
 PI
 XX
 XX WPI; 2000-491161/43.
 DR N-PSDB; AAA73293.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and

PT compound foods and feeds -
 XX Disclosure; Fig 25a-c; 240pp; English.
 PS
 XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 XX invention.
 SQ Sequence 467 AA;

Query Match 90.2%; Score 2093; DB 21; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLMGQYSPFFSLADESAISPDPVKGCRVTFVOVLSRHGARY 58
 DB 27 nshscdvtvgyqcfpeishlwgtyspffsladesaispdpvkgrvtfvqvlshrghary 86
 QY 59 PTSSSKKYSALIERIOKNAT-FKGKYAFKTYNTLGADDLTPFGENOMVNSGKIFYRR 117
 DB 87 ptssaskaysalieraiknatafkgkyafktyntlgadddtpfgeqmqvnsqikfyrr 146
 QY 118 YKALARNIVPFVRASGSDRVIASAEKFTIEGFSQAKLADPA---HQASPVINVIIEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaeekfiegfsqakladpaganphqaspinviiepegagy 206
 QY 175 NNTLDHGLCTAFEDSTLGDDEANFTAVFAPPPIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafeeseldgdeanftavfappirarleahpgvnltdedvnlmdmcpf 266
 QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatsqspfcldlftdewiqydyqlsgkygygagnpbgagvgvfnelliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDMTMSIFFALGLYNGTKPLSTTSV 350
 DB 327 lthspvqdhstnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsv 386
 QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEGEKEPLVRLVNDVRVPLHGGCVD 409
 DB 387 esieetdgyaswtvpfaarayvemmqcea-----ekeplvrvlndrvrvplhgcgvd 439
 QY 410 KLGRCKLDDFVEGLSFARSGGNNAECFA 437
 DB 440 klgrckrddfveglstarsgnweecfa 467

RESULT 10
 AAY69569
 ID AAY69569 standard; Protein; 467 AA.
 XX AAY69569;
 AC
 XX 19-APR-2000 (first entry)
 DT
 XX
 XX Mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.
 DE
 XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutin.
 XX Aspergillus terreus 9A1.
 OS

OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger str. NRRL3135.
 OS Aspergillus fumigatus ATCC13073.
 OS Aspergillus fumigatus ATCC32722.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Aspergillus nidulans.
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Paxillus involutus NN005693.
 OS Trametes pubescens NN9343.
 OS Agrocybe pediades NN009289.
 OS Peniophora lycii NN006113.
 OS Thermomyces lanuginosa.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..26 /note= "Phytase signal peptide from Aspergillus terreus
 FT cbs16.46"
 FT
 FT Protein 27..467
 FT /note= "Mature phytase-10-thermo[3]-Q50T-K91A"
 FT
 XX
 XX EP969089-A1.
 XX 05-JAN-2000.
 XX
 XX 23-JUN-1999; 99EP-0111949.
 XX
 XX 29-JUN-1998; 98EP-0111960.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 XX Brugger R, Lehmann M, Wyss M;
 PI WPI; 2000-099429/09.
 DR
 XX New stabilized enzyme formulation, useful for feed compositions for
 XX monogastric animals -
 XX Example '5; Fig 19; 101pp; English.
 XX
 CC The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-10
 CC consensus sequence, phytase-10-thermo[3]-Q50T-K91A, which has a
 CC temperature optimum and melting point 4 degrees Celsius higher than that
 CC of phytase-10 (AAY69566). Its specific activity with phytate as a
 CC substrate is also strongly increased.
 XX
 XX Sequence 467 AA;

Query Match 90.2%; Score 2093; DB 21; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;
 QY 1 NNSCDTVD-GYQC-PEISHLWGQYSPFFSLADESAISPDVPGKGRVTFVQVLSRGARY 58
 DB 27 nshscdtvdggygcfcpeishlwgtyspffsladesaispdvpgkgrvtfvqvlshrgary 86
 QY 59 PRSSKSKYSALIERIQKAT-FKGYAFLKTYNTLTGADDLTPFEGNQMVNSGKIFYYR 117
 DB 87 pssaskaysalieraiknatafkkyafktyntltgaddltfpgeqmvnsgikfyyr 146
 QY 118 YKALARNIVPFVRASGSDRVIASAEKIFEGFOSAKIADPA---HOASPVINVIIEGSGY 174
 DB 147 ykalarkiivpirasgsdrviasaekfiegfsakladpaganphqaspvinvilpegay 206
 QY 175 NNTLDHGLCTAFEDSTLGDDEANFTAVFAPPPIRARELEA-LPGVNLTDDEVDVNLMDMCP 233
 DB 207 nntldhglctafeeseldgdeanftavfappirareleahlpvnltdedvvnldmcp 266
 QY 234 DTVARTSDATOLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatqlspfcldlftdewiqydyqlsgkygygagngplgagvgvfnelliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTWVSFFALGLYNGTKPLSTTSV 350
 DB 327 lthspvqdhstnhtldsnpatplnatlyadfshtnmtvsiffalglyngtkplsttsv 386
 QY 351 ESI-ETDGYAASWTVPFAARAYVEMOCEAGGGEGEKEPLRVLRVNDRVVPLHGGVD 409
 DB 387 esieetdgyaswtvpfaarayvemqcea-----ekeplrvlrvndrvvplhggvd 439
 QY 410 KLGRCKLDDFVEGLSFARSGGNMACEFA 437
 DB 440 klgrckrddfveglslfarsggnweecfa 467
 RESULT 11
 AAB20531
 ID AAB20531 standard; Protein: 467 AA.
 XX
 AC AAB20531;
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX
 OS Synthetic.
 XX
 PN WO2000043503-A1.
 XX
 PD 27-JUL-2000.
 XX
 PF 21-JAN-2000; 2000WO-DK000025.
 XX
 PR 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX
 PA (NOVO) NOVO NORDISK AS.
 XX
 XX Lehmann M;
 XX
 XX WPI; 2000-491161/43.
 DR N-PSDB; AAY73290.
 DR
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and

PT compound foods and feeds -

PS Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g. file,
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 89.1%; Score 2067; DB 21; Length 467;
 Best Local Similarity 91.3%; Pred. No. 1.7e-193;
 Matches 409; Conservative 7; Mismatches 14; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PRISHLWQYSPFFSLADESAISPDPKGRVTFVQVLSRHGARY 58
 Db 27 nshscdvtvdygqpcpeishlwgtyspyfsladesaispdpkdcrtvfvqlsrhgary 86
 QY 59 PTKSKKYSALIERIOKNAT-FKGKYAFILKTYNTLGADLTTPFGENQWNSGKIFRYR 117
 Db 87 ptkskkysalieraiknatafkgyafiktyntlgadlttfgqndqmvnsgikfyrr 146
 QY 118 YKALARNIVFVRASGSDRVIAAEKFIQFQSAKLADPA--HQASPVINVIIPGSGY 174
 Db 147 ykalarivpfirasgsdrviasaekfiqfqsakladpqsqhqaspinviipgsgy 206
 QY 175 NNTLDHGLCTAFEDSTLGDDAENFTAVFAPPIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 Db 207 nntldhglctafedstlgsdveanfalfapalarlealdpgvntldedvylmdmcpf 266
 QY 234 DTVARTSDATQSPFCDLTADW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELLAR 290
 Db 267 dtvartsdatspfcalthdewiqdyqlsgkygygagndp1gpaqvgvfanellar 326
 QY 291 LTHSPVQDHTSNHILDSNPATFPLNATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSV 350
 Db 327 lthspvqdnstnhtldsnpatfplnatlyadfshdntmisiffalglyngtkplsttsv 386
 QY 351 EST-ETDGYAASWTVPFAARAYVENMOCCEAGGGGEGEKEPLVRVLVNDRVVPLHGCQVD 409
 Db 387 esleetdgyaswtvpfaarayvenmqcqa-----ekeplvrvlvndrvvplhgcavd 439
 QY 410 KLGRCKLDFVEGLSFARSGGNWAECA 437
 Db 440 klgrckrddfveglsfarsggnwaecca 467

RESULT 12

AAB20536
 ID AAB20536 standard; Protein; 424 AA.

XX AAB20536;

AC AAB20536;

XX 05-DEC-2000 (first entry)

DE Consensus phytase protein from Fig 4.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO2000043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 4a-d; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 424 AA;

Query Match 88.9%; Score 2064.5; DB 21; Length 424;
 Best Local Similarity 94.3%; Pred. No. 2.5e-193;
 Matches 414; Conservative 0; Mismatches 8; Indels 17; Gaps 12;

QY 1 NSHSCDVID-GYQCPEISHLWQYSPFFSLADESAISPDPKGRVTFVQVLSRHGARY 59
 Db 1 nshscdvtvdygqpcpeishlwgtyspyfsladesaispdpkdcrtvfvqlsrhgary 59
 QY 60 TSSKSKYSALIERIOKNATFKGYAFILKTYNTLGADLTTPFGENQWNSGKIFRYRYK 119
 Db 60 tsskskysali-aiknatfkgyafiktyntlgadlttfgedqgmvnsgikfyrryk 118
 QY 120 ALARNIVFVRASGSDRVIAAEKFIQFQSAKLADPAHQASPVINVIIPGSGYNTLTD 179
 Db 119 alar-ivpfvrasgsdrviasaekfiqfqsakladpa-qaspinviipg-gyntld 175
 QY 180 HGLCTAFEDSTLGDDAENFTAVFAPPIRARLEALPGVNLTDDEVVNLMDMCPDFVART 239
 Db 176 hgactafeseldgddveanfapppiralarlealpgvntldedvnlmdmcpdfvart 234
 QY 240 SDATQSPFCDLTADWQYDYLQSLKYGYGAGNPLGPAQGVGFNELLARLTHSPVQDH 299
 Db 235 sdattqlspfcldt-hewqdyqlsgkygygagndp1gpaqvgvfnellarlthspvqdh 293
 QY 300 TSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSVE-SIETDGY 358
 Db 294 tstnhtldsnpatfplnatlyadfshdntmvsiffalglyngt-plsttsvepseetdgy 352
 QY 359 AASWTVPFAARAYVENMOCCEAGGGGEGEKEPLVRVLVNDRVVPLHGCQVDKLGRCKLDD 418
 Db 353 aaswtvpfaarayvenmqce-----egekeplvrvlvndrvvplhgqcvdklgrck-dd 405

CC enzyme. The phytase formulation of the invention has an improved
CC thermostability and can therefore remain stable during long-term storage
CC and can withstand feed processing methods such as extrusion, expansion
CC and pelleting. The present sequence represents an initial consensus
CC phytase sequence determined using the program PRETTY from mature phytase
CC sequences from a variety of fungi (AAV69544-Y69546, AAV69548-Y69556,
CC AAV69564) and the Basidiomycetes phytase consensus AAV69563. This was
CC used as the basis for the consensus phytase designated phytase-10
CC (AAV69566).
XX
XX

Sequence 424 AA;

Query Match 88.9%; Score 2064.5; DB 21; Length 424;
Best Local Similarity 94.3%; Pred. No. 2.5e-193;
Matches 414; Conservative 0; Mismatches 8; Indels 17; Gaps 12;

Qy 1 NSHSCDVTVD-GYCPETISHLWGQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYP 59
Db 1 nshscdtvdggyqcpeishlwgygspffsladesalspdvp-gcrvtfvqvlsrhgaryp 59
Qy 60 TSSSKKYSALIERIQKNATFKGYAFLKTYNTLGGADDLTPFGENOMVNSGKFKVRYRK 119
Db 60 tssskkysali-aiqknatfkgyatfktyntlclgaddltlpgqeqmwnsgikfyrky 118
Qy 120 ALARNIVPFRASGDRVIAAEKFIQFQSAKLADPAHQASPVINVIIPESGSGYNNTLD 179
Db 119 alar-ivpfrasgsdrviasaekfiqfqsakladpa-qaspvlnvlipeg-gynntld 175
Qy 180 HGLCTAPEDSTLGGDDEANFTAVFAPPFIRARLEALPCVNLTDDEVDVNLMDMCPFDVART 239
Db 176 hglctafepseldgdeanftavfappfirarlealpcvnltdedvnlmdmcpfdvart 234
Qy 240 SDATOLSPFCDLFTADEWOYDQLQSLKYYGYGAGNPLGPAQGGVFNELIARLTHSPVDH 299
Db 235 sdatslspfcldft-hewdydqlslkyygygagnpplgpaqggvfnellarlthspvqdh 293
Qy 300 TSTNHTLDSNPATPLNATLYADESHDNTWVSIFPALGLYNGTKPLSTSVSVE-SIETDGY 358
Db 294 tstnhtldsnpatplnatlyadshdntwvsiffaiglyngt-plsttsvepseedgy 352
Qy 359 AASWTVPFAARAYVEMOCEAGGGGEGEKEPLVRVLVNDVRVPLHGGCVYDKLGRCKLDD 418
Db 353 aaswtvpfaarayvemmqce-----egekeplvrvlvndrvrplhbgcvdklgrck-dd 405
Qy 419 FVEGLSFARSGGNWAECEFA 437
Db 406 fveglsfarsggnweecfa 424

RESULT 14

AA205332
ID AAB20532 standard; Protein; 467 AA.

XX
AC AAB20532;

XX 05-DEC-2000 (first entry)

Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.

Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
temperature stability; pH profile; temperature profile; reaction rate;
specific activity; substrate specificity; substrate cleavage pattern;
substrate binding; position specificity; phytate degradation rate;
food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

PF

Qy 419 FVEGLSFARSGGNWAECEFA 437
Db 406 fveglsfarsggnweecfa 424
RESULT 13
ID AAV69565 standard; protein; 424 AA.
XX
XX AAV69565;
XX
XX 19-APR-2000 (first entry)
XX
XX Initial consensus mature phytase #2.

Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
thermostable; animal feed; monogastric animal; phytate phosphorus;
phosphate availability; consensus.

Aspergillus terreus 9A1.
Aspergillus terreus cbs16.46.
Aspergillus niger var. awamori.
Aspergillus niger str. NRRL3135.
Aspergillus fumigatus AFCC13073.
Aspergillus fumigatus ATCC32722.
Aspergillus fumigatus ATCC58128.
Aspergillus fumigatus ATCC26906.
Aspergillus fumigatus ATCC32239.
Emicella nidulans.
Talaromyces thermophilus ATCC20186.
Myceliophthora thermophila.
Paxillus involutus NN005693.
Trametes pubescens NN9343.
Agrocybe pediades NN009289.
Peniophora lycii NN006113.
Thermomyces lanuginosa.

XX EF969089-A1.

XX 05-JAN-2000.

XX 23-JUN-1999; 99EP-0111949.

XX 29-JUN-1998; 98EP-0111960.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Brugger R, Lehmann M, Wyss M;

XX WPI; 2000-099429/09.

XX New stabilized enzyme formulation, useful for feed compositions for
XX monogastric animals -

XX Example 4; Fig 16; 101pp; English.

The invention relates to a novel stabilised dry or liquid enzyme
formulation, comprising phytase (myo-inositol hexakisphosphate
phosphohydrolase) and one or more stabilising agents including
xyliol or ribitol; polyethylene glycols with a molecular weight of 600
to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
The stabilised phytase formulation is used in a method for preparing a
feed composition for monogastric animals (e.g., pigs, poultry) and
provides a monogastric animal with its dietary requirements of
phosphorus. Although a large amount of phosphate is present in animal
feed in the form of phytate phosphorus, monogastric animals are unable
to utilise this form of phosphate, resulting in the addition of extra
phosphate to the feed of such animals. Phytase enhances the nutritional
value of plant material without the need for adding additional phosphate
to the feed. The level of phosphate pollution in the environment is
reduced by adding phytase to animal feed, as the animal can make use of
the inorganic phosphate liberated from phytate phosphorus using the

XX PR 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX PA (NOVO) NOVO NORDISK AS.
 XX PI Lehmann M;
 XX XX WPI; 2000-491161/43.
 DR N-ESDB; AAA73291.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX XX Disclosure; Fig 23a-c; 240pp; English.
 PS XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX XX Sequence 467 AA;

Query Match 88.8%; Score 2061; DB 21; Length 467;
 Best Local Similarity 91.1%; Pred. No. 6.5e-193;
 Matches 408; Conservative 7; Mismatches 15; Indels 18; Gaps 10;
 QY 1 NSHSCDTPD-GYOC-PEISHLWQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARY 58
 DB 27 nshscdtdvggyqcfpeishlwgtyfpyfsladesaispdvpgkdrvtfvqlsthrgary 86
 QY 59 PTSSKSKYSALIERIQKNAT-FKGYAFKTYNTLGGADDLTPFGENQMVNSGKIFYRR 117
 DB 87 ptssaskaysalieleaiaqnatafkgyafktyntlgaddltpfgenqmvnsgikfyrr 146
 QY 118 YKALARNIVPFVRASGSDRVIASAEKFIQGSQAKLADPA---HOASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsvdrviasaekfiegfsqakladpgsqhqsasvpinvipegsgy 206
 QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPFIRARLEA-LPGVNLTDDEVVNLMDCPF 233
 DB 207 nntldhglctafedstlgdvveanftalfapairleadlpgvntldedvvyldmcpf 266
 QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatelspfcalfthdewiqdyqlsglkygygagngnplgpaqgvgfaneliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFPAGLYNGTKPLSTTSV 350
 DB 327 lthspvqdhstnhtldsnpatfplnatlyadfsdntmvsifpalyngtkplsttsv 386
 QY 351 ESI-ETDGYASWTVFAARAYVEMMOCEAGGGGEGEGEKEPLVRLVNDRVVPLHGCGVD 409
 DB 387 esietdgyaswstvfaarayvemmqcqa-----ekeplvrlvndrvvplhgcvd 439
 QY 410 KLGRCKLDDFVEGLSFARSNGNWAECFA 437
 DB 440 klgrckrdddfveglsfarsngnwaecfa 467

RESULT 15
 AAY43169

ID AAY43169 standard; Protein; 467 AA.
 XX AC AAY43169;
 XX XX 06-JAN-2000 (first entry)
 DT XX
 DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.
 XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 KW consensus sequence.
 XX OS Synthetic.
 XX XX WO9948380-A1
 PN PD 30-SEP-1999.
 XX XX 22-MAR-1999; 99WO-DK00154.
 PF XX 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 PR 22-JAN-1999; 99DK-0000093.
 XX (NOVO) NOVO-NORDISK AS.
 PA Petersen S;
 XX WPI; 1999-591030/50.
 DR N-ESDB; AAZ31520.
 XX XX Preparing animal feed using a thermostable phytase -
 PT Example 3; Fig 9; 71pp; English.
 PS This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.
 XX SQ Sequence 467 AA;

Query Match 88.1%; Score 2044; DB 20; Length 467;
 Best Local Similarity 90.4%; Pred. No. 3e-191;
 Matches 405; Conservative 7; Mismatches 18; Indels 18; Gaps 10;
 QY 1 NSHSCDTPD-GYOC-PEISHLWQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARY 58
 DB 27 nshscdtdvggyqcfpeishlwgtyfpyfsladesaispdvpgkdrvtfvqlsthrgary 86
 QY 59 PTSSKSKYSALIERIQKNAT-FKGYAFKTYNTLGGADDLTPFGENQMVNSGKIFYRR 117
 DB 87 ptssaskaysalieleaiaqnatafkgyafktyntlgaddltpfgenqmvnsgikfyrr 146
 QY 118 YKALARNIVPFVRASGSDRVIASAEKFIQGSQAKLADPA---HOASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsvdrviasaekfiegfsqakladpgsqhqsasvpinvipegsgy 206
 QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPFIRARLEA-LPGVNLTDDEVVNLMDCPF 233
 DB 207 nntldhglctafedstlgdvveanftalfapairleadlpgvntldedvvyldmcpf 266
 QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatelspfcalfthdewiqdyqlsglkygygagngnplgpaqgvgfaneliar 326

Sat Oct 27 15:25:41 2001

QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFFALGLYNGTKPLSTTSV 350
Db 327 lthspvqdhstnhtlidsnpatfplnatlyadfsdntmisiffalglyngtkplsttsv 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRLVNDRVVPLHGCQVD 409
Db 387 esieetdgysaswtvpfaarayvemmqcqa-----ekeplrvivndrvvpplhgcavd 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 klgrckrddfrveglsfarsggnwaecefa 467

Search completed: October 26, 2001, 16:40:04
Job time: 4955 sec

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2020	87.0	441	4	US-09-121-425-1	Sequence 1, Appli	
2	1902	81.9	467	4	US-09-121-425-2	Sequence 2, Appli	
3	1698	72.7	467	1	US-07-923-724-8	Sequence 8, Appli	
4	1688	72.7	467	2	US-08-609-426A-8	Sequence 8, Appli	
5	1698	72.7	467	2	US-08-374-652C-2	Sequence 2, Appli	
6	1684	72.6	467	1	US-08-151-574-32	Sequence 32, Appli	
7	1684	72.6	467	1	US-08-146-424-20	Sequence 20, Appli	
8	1684	72.6	467	1	US-08-693-709-2	Sequence 2, Appli	
9	1684	72.6	467	2	US-08-419-448-32	Sequence 32, Appli	
10	1684	72.6	467	2	US-08-819-825-3	Sequence 3, Appli	
11	1684	72.6	467	4	US-09-163-642-3	Sequence 3, Appli	
12	1672	72.0	443	4	US-02-155-855-1	Sequence 1, Appli	
13	1672	72.0	467	4	US-09-155-855-3	Sequence 3, Appli	
14	1667	71.8	443	4	US-09-155-855-2	Sequence 2, Appli	
15	1322	57.0	475	2	US-08-819-825-2	Sequence 2, Appli	
16	1322	57.0	475	4	US-09-163-642-2	Sequence 2, Appli	
17	837.5	36.1	443	3	US-08-993-359-30	Sequence 30, Appli	
18	812.5	35.0	453	3	US-08-993-359-22	Sequence 22, Appli	
19	804.5	34.7	439	3	US-08-993-359-24	Sequence 24, Appli	
20	804.5	34.7	439	3	US-09-321-654-2	Sequence 2, Appli	
21	804.5	34.7	439	3	US-08-969-359A-2	Sequence 2, Appli	
22	793	34.2	442	3	US-08-993-359-28	Sequence 28, Appli	
23	791.5	34.1	442	3	US-08-993-359-26	Sequence 26, Appli	
24	343.5	14.8	446	1	US-07-627-539G-7	Sequence 7, Appli	
25	343.5	14.8	468	1	US-07-627-539G-2	Sequence 2, Appli	
26	307.5	13.2	479	1	US-07-923-724-2	Sequence 2, Appli	
27	307.5	13.2	479	2	US-08-609-426A-2	Sequence 2, Appli	

Db 301 LTRSPVQDHTSTNHTLDSNPATPLNATLYADFSHDNSMISIPFALGLVNGTAPLSTTSV 360
Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRLVNDVRVPLHOCGYD 409
Db 361 ESIETDGYASWTVPFARAYVEMMQCEA-----EKEPLVRLVNDVRVPLHGCAYD 413
Qy 410 KLGRCKLDDFVEGLSFARSGGNNAECFA 437
Db 414 KLGRCKRDDFVEGLSFARSGGNNAECFA 441
RESULT 2
US-09-121-425-2
; Sequence 2, Application us/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: sequence
US-09-121-425-2
Query Match 81.9%; Score 1902; DB 4; Length 467;
Best Local Similarity 81.8%; Pred. No. 2.9e-195;
Matches 382; Conservative 10; Mismatches 19; Indels 56; Gaps 11;
Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NSHSCDVTVDGQYQCFPEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 86
Qy 59 PTSSKSKYSALIERIQKN-ATEFGYAFKLTNYTLGADDLTPFGENQVNSGKIFRYR 118
Db 87 PTSSKSKYSA-----TNYTLGADDLTPFGENQVNSGKIFRYR 127
Qy 119 KALARNIVPFRASGSDRVIASAEKFIQFQSAKLADPA---HQASPVIN----- 165
Db 128 KALARKIVPFRASGSDRVIASAEKFIQFQSAKLADPGSQPHQASPVIDLIEATQKNAT 187
Qy 166 -----VPIEGSGYNNYLDHGLCTAFEDSTLGDAAEANTFATVPPIRABLEA-L 214
Db 188 AFRGKYAFLKVIPEGSGYNNYLDHGLCTAFEDSTLGDAAEANTFATVPPIRABLEADL 247
Qy 215 PGVNLTDVNNLMDMCPDFTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYIGYGA 272
Db 248 PGVNLTDVNNLMDMCPDFTVARTSDATQSPFCALFTHDEMROYDYLSGLKYYGYGA 307
Qy 273 GNPLGPAQGVGF-NELIARLTHSPVDHTSTNHTLDSNPATPLNATLYADFSHDNTWVS 331
Db 308 GNPLGPAQGVGFANELIARLTHSPVDHTSTNHTLDSNPATPLNATLYADFSHDNTWVS 367
Qy 332 IFPAGLNGTKPLSTSVESI-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEP 390
Db 368 IFPAGLNGTKPLSTSVESIETDGYASWTVPFARAYVEMMQCEA-----EKEP 420
Qy 391 LVRVNDVRVPLHGGGVCKLGRCKLDDFVEGLSFARSGGNNAECFA 437
Db 421 LVRVNDVRVPLHGGGVCKLGRCKRDDFVEGLSFARSGGNNAECFA 467
RESULT 3

US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8
Query Match 72.7%; Score 1688; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 2.6e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;
Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NOSTCDVTVDQYQCFSTSHLWGOYAPFFSLANESALSPDVPGKCRVTFVQVLSRHGARY 86
Qy 59 PTSSKSKYSALIERIQKN-ATEFGYAFKLTNYTLGADDLTPFGENQVNSGKIFRYR 117
Db 87 PTESKGGKYSALIEIQNVTFDQKYAFKLTNYSLGADDLTPFGQELVNSGKIFQYR 146
Qy 118 YKALARNIVPFRASGSDRVIASAEKFIQFQSAKLADP---HQASPVINVIPEGSGY 174
Db 147 YESLTRNIPFIRSSGSRVIAASGEKFIQFQSTKLKDPRAQPGQSGPKIDVVISEASS 206

```

QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFAVAPPFIRABLE-ALPGVNLTDDEVDVNLMDMCPF 233
Db 207 NNTLDPGCTCTVEDESELADTVEANFTAFAPSIQRLENLDSGVTLTDTETVYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTDEWHIDYLOSLKKYGGAGNPLGPTQGVGVANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHONTWYSIFFALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSHONGIISILFALGLYNGTKPLSTTTV 386
QY 351 EST-ETDGYAASWTVPFAARYVEMMOCEAGGGGEGEKEPLVRVLVNDVRVPLHSCGVD 409
Db 387 ENITQDTDFSSAWTVPFASRLYVEMMQCA-----EQEPLVRVLVNDVRVPLHSCPID 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECSA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 4
US-08-609-426A-8
; Sequence 8, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerström, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; INVENTOR: in Trichoderma
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.
; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600

```

```

; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-426A-8

```

```

Query Match 72.7%; Score 1688; DB 2; Length 467;
Best Local Similarity 73.9%; Pred. No. 2.6e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;
QY 1 NNSHCDTVD-GYQC-PEJSHLWGOVSPFSLADESAISPDVPKGRVTFVVLRSRHGARY 58
Db 27 NOSTCDTVDGQYQCFSETSHLWGOVAPFFSLANESAISSPDVPAGCRVTFVVLRSRHGARY 86
QY 59 PTSSKSKYSALIERIQKN-ATFKGYAFLKTYNTVTLGADDLTPFGENOMVNSGIKEYRR 117
Db 87 PTESKGGKYSALIEIQONVTFDGYAFLKTYNSLGADDLTFGQELVNSGIKEYRR 146
QY 118 YKALARNIVPFRASGSDRVFTASAEKFTIEGQSALADP---AHQASPVINVIIEGSGY 174
Db 147 YESLTRNIIPFIRSSGSRVFTASAEKFTIEGQSALADP---AHQASPVINVIIEGSGY 206
QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFAVAPPFIRABLE-ALPGVNLTDDEVDVNLMDMCPF 233
Db 207 NNTLDPGCTCTVEDESELADTVEANFTAFAPSIQRLENLDSGVTLTDTETVYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTDEWHIDYLOSLKKYGGAGNPLGPTQGVGVANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHONTWYSIFFALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSHONGIISILFALGLYNGTKPLSTTTV 386
QY 351 EST-ETDGYAASWTVPFAARYVEMMOCEAGGGGEGEKEPLVRVLVNDVRVPLHSCGVD 409
Db 387 ENITQDTDFSSAWTVPFASRLYVEMMQCA-----EQEPLVRVLVNDVRVPLHSCPID 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECSA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467

```

```

RESULT 5
US-08-374-652C-2
; Sequence 2, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

```

10

Query Match	72.6%;	Score 1684;	DB 1;	Length 467;
Best Local Similarity	73.9%;	Pred. No. 6.9e-172;		
Matches 331; Conservative	40;	Mismatches 59;	Indels 18;	Gaps
1	NSHSCDVID-GYQC-PEISHLWQYSPFSLADESAISPDPKGRVTFVQVLSRHGARY	58		
QY				
27	NOSSCDTVDGYQCFSETSHLWQYAPFFSLANESVISPFAVGRVTFVQVLSRHGARY	86		
DB				
59	PTSSKSKKYSALIERIOKNA-TFKGKYAFUKTYNTLGADDTLPFGENOMWNSGKIFVRR	117		
QY				
87	PTDSKSKKYSALIEEQONATFDGKYAFUKTYNSLGDADDTLPFGEQELVNSGKIFQR	146		
DB				
118	YKALARNIVPEVRASGSDRVIASAEKFIEGFQSAKLADP---AHOQASPVINVIIEPGSY	174		
QY				
147	YESLTRNIVPPIRSGSSRVISACKFIEGFQSTKLDPRAOPQSSPIDVIVISEASS	206		
DB				
175	NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPFARLARL-ALPGVNLTDEDEVNLMDCPF	233		
QY				
207	NNTLDPGFCTVFDESELAADTVEANFTATFVPSIRQRLNDLSGVTLTDTEVYLMDCSF	266		
DB				
234	DTVAARTSDATQLSPFCDLTFADEW-QYDYLOSL-KYYGYGAGNPLGPAQGVF-NELIAR	290		
QY				
267	DTISTSTVDTKLSPFCDLFTHDEWINDYLOSLKYYGYGHGAGNPLGPTQGVGYANELIAR	326		
DB				
291	LPHSPVQDHTSTNHTLDSNPATPFLNATLYADSHDNTWYSIFFALGLYNGTKPLSTTSV	350		
QY				
327	LPHSPVHDTSNHTLDOSSPATPFLNLTLYADSHDNGIITISILFALGLYNTKPLSTTIV	386		
DB				

10

```

QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEKEPLVRLVNDVRVPLHGCGVD 409
Db 387 ENITQDGFSSAWTVPFASRLYVEMMOCEA-----EQEPLVRLVNDVRVPLHGCPVD 439
QY 410 KLGRCCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDTSFVRGLSFARSGGNWAECEFA 467

RESULT 7
US-08-146-424-20
; Sequence 20, Application US/08146424
; Patent No. 5593963
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,424
; FILING DATE: 02-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KENNEDY, BILL
; REGISTRATION NUMBER: 33,407
; REFERENCE/DOCKET NUMBER: 44615-20011.24
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-146-424-20

Query Match 72.6%; Score 1684; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKGVTFVQVLSRHGARY 58
Db 27 NOSSCDTVDQGYQCFSETSHLWGOYAPFFSLANESVISPVPACRVTFAQVLSRHGARY 86
QY 59 PTSSKSKYSALIEIQKNA-TFGKVAFLKTYNTTGADLTPFGENQMVNSGKIFYRR 117
Db 87 PTDSKGKYSALIEIQNAFTDPCKVAFLKTYNSLGADLTPFGQELVNSGKIFYQR 146
QY 118 YKALARNIVPVRAGSGDRVIAAEKTEGFQSAKLADP---AHQASPVNVIPEGSGY 174
Db 147 YESLNRNIVPIRSGSGSRVIAASKKTEGFQSKLADPRAQPGQSSPKIDVWVISEASS 206
QY 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPTRARLE-ALPGVNLTDDEVDWNLMDMCPF 233

```

```

Db 207 NNTLDPGTCTVEDSELADTVEANFTATVPSPTRQLENDLSGVLTDEVTYLMDMCSF 266
QY 234 DTVARTSDATOLSPDCDLETADGW-QYDYLQSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPDCDLETFHDEWINYDYLQSLKYYGYGAGNPLGPTQGVGYANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFFALGYNKTKPLSTTSTV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFSDHNGIISILFALGYNKTKPLSTTIV 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEKEPLVRLVNDVRVPLHGCGVD 409
Db 387 ENITQDGFSSAWTVPFASRLYVEMMOCEA-----EQEPLVRLVNDVRVPLHGCPVD 439
QY 410 KLGRCCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDTSFVRGLSFARSGGNWAECEFA 467

RESULT 8
US-08-693-709-2
; Sequence 2, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1...23
; OTHER INFORMATION:

```

US-08-693-709-2

Query Match 72.6%; Score 1684; DB 1; Length 467;

Best Local Similarity 73.9%; Pred. No. 6.9e-172; Indels 18; Gaps 10;
Matches 331; Conservative 40; Mismatches 59;

```
QY 1 NSHSCDVTVD-GYOC-PEISHLMGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSSCDTVDQGYQCFSETSHLMGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIOKNA-TFKGKYAFKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117
Db 87 PTDSKGRKYSALIEIOONATTFDGKYAFKTYNSLGGADDLTPFGEOLVNSGKIFYYR 146
QY 118 YKALARNIVPVRASGSDRVIASAEKFTIEGFQSAKLADP---AHQASPVINVIPEGSY 174
Db 147 YESLTRNIVPFISSGSRVSIASGKFTIEGFQSTKLDPRAGQSSPKIDVIVSEASSS 206
QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVAPPPIRABLE-ALPGVNLTDDEVVNLMDMCPF 233
Db 207 NNTLDPGCTVFEDSELADTVEANFTATFVPSIRQRLNDSGLVTLDTTEVTYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPPCDLFTHDEWINYDYLQSLKKYKHGAGNPLGTPQGVGYANELLAR 326
QY 291 LTHSPVQDHTSNHTPLDSDNPATFPLNATLYADFSHDNTMWSIFFALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLSSPATFPLNSTLYADFSHDNGIIISILFALGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGEKEPLRVRLVNDRVVPLHGCYVD 409
Db 387 ENITQDTGFSASWTVPFASRLYVEMMQCOA-----EQEPLRVRLVNDRVVPLHGCYVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECEFA 467
```

RESULT 9

US-08-419-448-32

; Sequence 32, Application US/08419448

; Patent No. 5863533

; GENERAL INFORMATION:

; APPLICANT: Robert F.M. Van Gorcom

; APPLICANT: Willem Van Hartingsveldt

; APPLICANT: Petrus A. Van Paridon

; APPLICANT: Annemarie E. Veenstra

; APPLICANT: Rudolf G.M. Luttin

; APPLICANT: Gerardus Seltin

; TITLE OF INVENTION: Cloning and Expression of Microbial

; TITLE OF INVENTION: Phytase

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster

; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20006-1888

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; SOFTWARE:

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/419,448

; FILING DATE: 10-APR-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Murashige, Kate H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 24615-20026.10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-887-1500

; INFORMATION FOR SEQ ID NO: 32:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 467 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-419-448-32

Query Match 72.6%; Score 1684; DB 2; Length 467;

Best Local Similarity 73.9%; Pred. No. 6.9e-172;

Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

```
QY 1 NSHSCDVTVD-GYOC-PEISHLMGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSSCDTVDQGYQCFSETSHLMGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIOKNA-TFKGKYAFKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117
Db 87 PTDSKGRKYSALIEEQNATTFDGKYAFKTYNSLGGADDLTPFGEOLVNSGKIFYYR 146
QY 118 YKALARNIVPVRASGSDRVIASAEKFTIEGFQSAKLADP---AHQASPVINVIPEGSY 174
Db 147 YESLTRNIVPFISSGSRVSIASGKFTIEGFQSTKLDPRAGQSSPKIDVIVSEASSS 206
QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVAPPPIRABLE-ALPGVNLTDDEVVNLMDMCPF 233
Db 207 NNTLDPGCTVFEDSELADTVEANFTATFVPSIRQRLNDSGLVTLDTTEVTYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPPCDLFTHDEWINYDYLQSLKKYKHGAGNPLGTPQGVGYANELLAR 326
QY 291 LTHSPVQDHTSNHTPLDSDNPATFPLNATLYADFSHDNTMWSIFFALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLSSPATFPLNSTLYADFSHDNGIIISILFALGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGEKEPLRVRLVNDRVVPLHGCYVD 409
Db 387 ENITQDTGFSASWTVPFASRLYVEMMQCOA-----EQEPLRVRLVNDRVVPLHGCYVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECEFA 467
```

RESULT 10

US-08-819-825-3

; Sequence 3, Application US/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berkka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/819,825
FILING DATE: 18-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4758, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-819-825-3

Query Match 72.6%; Score 1684; DB 2; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PEISHLWQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSSCDTVDQGYQCFSESHLWQYAPFFSLANESVISPVPAGCRVTTAQVLSRHGARY 86
QY 59 PTKSKKYSALIERIQKNA-TFKGYAFLKTYNTLGADDLTPFGENQWNSGKIFYRR 117
Db 87 PTKSKKYSALIEEQONATTFDGKYAFLKTYNTSLGADDLTPFGELVNSGKIFYQR 146
QY 118 YKALARNIVPVRASGSDRVIAAEKIEGFSQAKLADP---AHQASPVNVIIPGSGY 174
Db 147 YESLTRNIVPIRSSGSRVIASSGKIEGFSQAKLADP---AHQASPVNVIIPGSGY 206
QY 175 NNTLDHGLCTAFEDSLTGDAAEANTFVAPPFIRARLE-ALPGVNLTDDEVNLMDCPF 233
Db 207 NNTLDPGTCTVFEDSELADTVEANFTATFVPSIRQLENDLSGVTLDTEVYLMDCSF 266
QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYQLSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSFPFCDLFTHDEWINDYQLSLKYYGYGAGNPLGPTQGVYANELIAR 326
QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASVTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDRVVPLHGCVD 409
Db 387 ENTQTQDGFSSAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDRVVPLHGCVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECA 467

RESULT 11
US-09-163-642-3
Sequence 3, Application US/09163642
Patent No. 6221644
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Ray, Michael W.
APPLICANT: Klotz, Alan V.
TITLE OF INVENTION: Polypeptides Having Phytase Activity
TITLE OF INVENTION: And Nucleic Acids Encoding Same
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6221644 No. 6221644disk of No. 6221644th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York

COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/819,825
FILING DATE: 18-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4758, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-163-642-3

Query Match 72.6%; Score 1684; DB 4; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PEISHLWQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSSCDTVDQGYQCFSESHLWQYAPFFSLANESVISPVPAGCRVTTAQVLSRHGARY 86
QY 59 PTKSKKYSALIERIQKNA-TFKGYAFLKTYNTLGADDLTPFGENQWNSGKIFYRR 117
Db 87 PTKSKKYSALIEEQONATTFDGKYAFLKTYNTSLGADDLTPFGELVNSGKIFYQR 146
QY 118 YKALARNIVPVRASGSDRVIAAEKIEGFSQAKLADP---AHQASPVNVIIPGSGY 174
Db 147 YESLTRNIVPIRSSGSRVIASSGKIEGFSQAKLADP---AHQASPVNVIIPGSGY 206
QY 175 NNTLDHGLCTAFEDSLTGDAAEANTFVAPPFIRARLE-ALPGVNLTDDEVNLMDCPF 233
Db 207 NNTLDPGTCTVFEDSELADTVEANFTATFVPSIRQLENDLSGVTLDTEVYLMDCSF 266
QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYQLSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSFPFCDLFTHDEWINDYQLSLKYYGYGAGNPLGPTQGVYANELIAR 326
QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASVTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDRVVPLHGCVD 409
Db 387 ENTQTQDGFSSAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDRVVPLHGCVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECA 467

RESULT 12
US-09-155-855-1
Sequence 1, Application US/09155855
Patent No. 6139902
GENERAL INFORMATION:
APPLICANT: KONDO, Hidemasa

; APPLICANT: ANAZAWA, Hideharu
 ; APPLICANT: KANEKO, Syunichi
 ; APPLICANT: NAGASHIMA, Tadashi
 ; APPLICANT: TANGE, Tatsuya
 ; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
 ; FILE REFERENCE: 81356/124
 ; CURRENT APPLICATION NUMBER: US/09/155,855
 ; EARLIER FILING DATE: 1998-10-05
 ; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175
 ; EARLIER FILING DATE: 1997-04-04
 ; EARLIER APPLICATION NUMBER: JP 084314
 ; EARLIER FILING DATE: 1996-04-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 443
 ; TYPE: PRT
 ; ORGANISM: Aspergillus niger
 ; US-09-155-855-1

Query Match 72.0%; Score 1672; DB 4; Length 443;
 Best Local Similarity 72.8%; Pred. No. 1.2e-170;
 Matches 326; Conservative 46; Mismatches 58; Indels 18; Gaps 10;
 QY 1 NSHSCDTVD-GYQC-PEISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
 Db 3 NOSTCTDVGQYCFSETSHLWGOYAPFFSLANKSAISPDVPAGCHVTFPAQVLSRHGARY 62
 QY 59 PTSSKSKYSALIERIQKNA-TFKGKYAFLKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117
 Db 63 PTDSKGGKYSALIEIQONATTFEGKYAFLKTYNYSLGADDLTPFGEQELVNSGVKFYQR 122
 QY 118 YKALARNIVPFVRASGSDRVASAEKFTFEGQSALADP---AHQASPVINVIIEGSGY 174
 Db 123 YESLTRNIVPFIRSSGSRVIAAGNKFIEGFQSTKLKDPRAQGOSSPKIDVIVISEASTS 182
 QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233
 Db 183 NNTLDPGCTVFEDELADDEANFTATFVPSIRQLENDLSGVSITDTEVTYLMDCSF 242
 QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 Db 243 DTISTSTVDTKLSPFCDLFTHEEWINDYLSLNKYHGGAGNPLGPTQGVGYANELIAR 302
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPALGLYNGTKPLSTTSV 350
 Db 303 LTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFSDNGIISILFALGLYNGTKPLSSTA 362
 QY 351 EST-ETDGYAASVTVPFAARAYVEMMQCEAGGGGEGEKEPLRVRLVNDVRVPLHGCQVD 409
 Db 363 ENITQIDGFSANTVPFAARMYVEMMQCS-----EQEPLRVRLVNDVRVPLHGCQVD 415
 QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 416 ALGRCTRDSEVFKLSFARSGGDWGECEFA 443

RESULT 13
 US-09-155-855-3
 ; Sequence 3, Application US/09155855
 ; Patent No. 6139902
 ; GENERAL INFORMATION:
 ; APPLICANT: KONDO, Hidemasa
 ; APPLICANT: ANAZAWA, Hideharu
 ; APPLICANT: KANEKO, Syunichi
 ; APPLICANT: NAGASHIMA, Tadashi
 ; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
 ; FILE REFERENCE: 81356/124
 ; CURRENT APPLICATION NUMBER: US/09/155,855
 ; EARLIER FILING DATE: 1998-10-05
 ; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175

; EARLIER FILING DATE: 1997-04-04
 ; EARLIER APPLICATION NUMBER: JP 084314
 ; EARLIER FILING DATE: 1996-04-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 467
 ; TYPE: PRT
 ; ORGANISM: Aspergillus niger
 ; US-09-155-855-3
 Query Match 72.0%; Score 1672; DB 4; Length 467;
 Best Local Similarity 72.8%; Pred. No. 1.3e-170;
 Matches 326; Conservative 46; Mismatches 58; Indels 18; Gaps 10;
 QY 1 NSHSCDTVD-GYQC-PEISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
 Db 27 NOSTCTDVGQYCFSETSHLWGOYAPFFSLANKSAISPDVPAGCHVTFPAQVLSRHGARY 86
 QY 59 PTSSKSKYSALIERIQKNA-TFKGKYAFLKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117
 Db 87 PTDSKGGKYSALIEIQONATTFEGKYAFLKTYNYSLGADDLTPFGEQELVNSGVKFYQR 146
 QY 118 YKALARNIVPFVRASGSDRVASAEKFTFEGQSALADP---AHQASPVINVIIEGSGY 174
 Db 147 YESLTRNIVPFIRSSGSRVIAAGNKFIEGFQSTKLKDPRAQGOSSPKIDVIVISEASTS 206
 QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233
 Db 207 NNTLDPGCTVFEDELADDEANFTATFVPSIRQLENDLSGVSITDTEVTYLMDCSF 266
 QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 Db 267 DTISTSTVDTKLSPFCDLFTHEEWINDYLSLNKYHGGAGNPLGPTQGVGYANELIAR 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPALGLYNGTKPLSTTSV 350
 Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFSDNGIISILFALGLYNGTKPLSSTA 386
 QY 351 EST-ETDGYAASVTVPFAARAYVEMMQCEAGGGGEGEKEPLRVRLVNDVRVPLHGCQVD 409
 Db 387 ENITQIDGFSANTVPFAARMYVEMMQCS-----EQEPLRVRLVNDVRVPLHGCQVD 439
 QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 440 ALGRCTRDSEVFKLSFARSGGDWGECEFA 467

RESULT 14
 US-09-155-855-2
 ; Sequence 2, Application US/09155855
 ; Patent No. 6139902
 ; GENERAL INFORMATION:
 ; APPLICANT: KONDO, Hidemasa
 ; APPLICANT: ANAZAWA, Hideharu
 ; APPLICANT: KANEKO, Syunichi
 ; APPLICANT: NAGASHIMA, Tadashi
 ; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
 ; FILE REFERENCE: 81356/124
 ; CURRENT APPLICATION NUMBER: US/09/155,855
 ; EARLIER FILING DATE: 1998-10-05
 ; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175
 ; EARLIER FILING DATE: 1997-04-04
 ; EARLIER APPLICATION NUMBER: JP 084314
 ; EARLIER FILING DATE: 1996-04-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 443
 ; TYPE: PRT
 ; ORGANISM: Aspergillus niger

Sat Oct 27 15:25:42 2001

us-09-488-265-27.ra1

Page 10

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db	51	CGGATGCCGACGATGGTGGTGGACGCTGACGAAAGTTGGTTGAGACGT	1
RESULT	2		
LOCUS	AQ255459	837 bp	DNA
DEFINITION	mgxb0014M15r	CGI Rice Blast BAC Library	Magnaporthe grisea genomic clone mgxb0014M15r, DNA sequence.
ACCESSION	AQ255459		
VERSION	AQ255459.1	GI:3779774	
KEYWORDS	GSS.		
SOURCE	Magnaporthe grisea.		
ORGANISM	Magnaporthe grisea		
REFERENCE			
AUTHORS	Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.		
TITLE	A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome		
JOURNAL	Unpublished (1998)		
COMMENT	Contact: Dean RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson University, Clemson, SC 29634 Tel: 864 656 5737 Fax: 864 656 4293 Email: rdean@clemson.edu Seq primer: GGAACAGCTATGACCATG Class: BAC ends High quality sequence start: 60 High quality sequence stop: 277. Location/Qualifiers		
FEATURES	source		
	1..837		
	/organism="Magnaporthe grisea"		
	/strain="70-15"		
	/db_xref="taxon:148305"		
	/clone="mgxb0014M15r"		
	/clone_lib="CGI Rice Blast BAC Library"		
	/tissue_type="Protoplasts"		
	/lab_host="E. coli DH10B"		
	/note="Vector: pBACWICH; Site1: HindIII; Site2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."		
BASE COUNT	248 a	154 c	188 g
ORIGIN			243 t
			4 others
Query Match	8.7%;	Score 122.4;	DB 226; Length 837;
Best Local Similarity	55.7%;	Pred. No. 2.9e-25;	
Matches	234; Conservative	0; Mismatches 186;	Indels 0; Gaps 0;
Qy	710	cagctattagatgagtggaagctgactgcccaggtttactttgactgacgagacg	769
Db	113	CCGGTCTTATACACGATAGGCTGGGATTTTCAGGTGCTGGGTTAAACGGTCAAGACG	172
Qy	770	ttgttactgagcatgtgtccattgacactgtcgtagaacttctgacgtactgt	829
Db	173	TTGTAATCTATGCGGACTGTGCACATGGAACACGCGCACTTTGAAACACTGGCC	232
Qy	830	aattgtctccattgtgtgttttctcactcacgacgaatggatccactgactttgc	889
Db	233	AACCTTCACCGCTTGGCAATCTGTTACGGAAGCAGATTGGGTAAATATGTTACTGT	292
Qy	890	aaagcttgggttaagtaaggtacggttagctgtgttaacccattgggtccagctcaagg	949
Db	293	CCAGCGTGCARAAATGGTACAGATACGGAACGGTAATCTTTGGGCCCACTATGGCG	352
Qy	950	ttggtttgctaacgaattgattgctagatgactcactctccagttcaagacacactt	1009
Db	353	TGGGATGGTAAACGAACCTTATTGCACGATTAAACCGGAAGCCAGTCCAAAGATCA	412
Qy	1010	ctactaacacacacttgactctaacccagctactttccattgaacgctactttgtacg	1069
Db	413	TGTCCAATACGACACTTGACATGAACCCGGAACCTTACTCTACAGGCAAAATTGTATG	472
Qy	1070	ctgacttctctcacgacacacactatgatctctatttttcttggcttttgggtttgtaca	1129
Db	473	CTGATTTTATCCATACGACGATATATAGGATTTAGCCGCTTTGGCTATTAAACG	532
RESULT	3		
LOCUS	AJ274007	581 bp	mRNA
DEFINITION	AJ274007 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae		
ACCESSION	AJ274007		
VERSION	AJ274007.1	GI:6433380	
KEYWORDS	EST.		
SOURCE	Metarhizium anisopliae.		
ORGANISM	Metarhizium anisopliae		
REFERENCE	1 (bases 1 to 581)		
AUTHORS	Screen, S.E., Mathur, P. and St. Leger, R.J.		
TITLE	EST analysis of the insect pathogenic fungus Metarhizium anisopliae		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Screen SE Entomology University of Maryland 4112 Plant Sciences Building, College Park, MD 20742, USA. Location/Qualifiers		
FEATURES	source		
	1..581		
	/organism="Metarhizium anisopliae"		
	/strain="ARSEF 2575"		
	/db_xref="taxon:5530"		
	/clone="Wa#1628"		
	/clone_lib="Metarhizium anisopliae ARSEF 2575"		
	/note="Vector: Unizap; Metarhizium anisopliae was grown on insect cuticle for 24 hours. A cDNA library was constructed in the unidirectional Lambda vector, Unizap"		
BASE COUNT	138 a	153 c	146 g
ORIGIN			144 t
Query Match	7.6%;	Score 107.4;	DB 104; Length 581;
Best Local Similarity	57.5%;	Pred. No. 7.7e-21;	
Matches	233; Conservative	0; Mismatches 166;	Indels 6; Gaps 2;
Qy	142	tgggtgacttacttcatacttcttcttggcagacgaatctgtctattttccagagctt	201
Db	163	TGGGCGGAGTACTCGCGTACTTCTC---AGCACACAAGGGATCTATTAAAGCCTGACATC	219
Qy	202	ccagacgacttagagtacttttctagagttttcttagacacggtgtagataccaca	261
Db	220	CCGTGCGGATCGAGGTCACCTTTGCTTCAATCTTGTCTCGACATGGGTCCCGAACCCG	279
Qy	262	acttcttctgcgttaaggcttactgtctttgtattgaagctatttcaaaagaacgtact	321
Db	280	ACGGCGGCAAGTCAAGGCTCAAGGACCTGGTTGAGCGGATCCAAAAGAGAGTCAAG	339
Qy	322	gctttcaagggttaagtagcgttttcttgaagacttacaactacacttttgggtgtagac	381
Db	340	GACTATGGAAGGGCTTCAGTTTCTCAAAGATTACAAGTATACTTAGGTCTTGAGCAC	399
Qy	382	ttgactccattcgtgtaaaacaaatggttaactctgtttaagttctacagagatcac	441
Db	400	CTCACCCCTTGGGGGAGCAGGAAATGGTCAAGTCTGGAAAAAAGTTCTTCAACGCTAT	459

QY	442	aaggctttggcgtaaa:--gattgttcattcattagagcttcttggtctgcacagatt	498
Dd	460	CAGAAGTTCGGGAGGATTTCGACCCATCCCTTTGTTCGAGCCTCGGCTCTGAAGAGTC	519
QY	499	atgtctctactgaagaagtccattgaaggtttccaatctgtctaag	543
Dd	520	GTCATGTCAGCGCAAGATTTGTTTCATGGCTTCTACAAGGCCAAG	564

RESULT	4
LOCUS	AQ163004
DEFINITION	AQ163004 753 bp DNA GSS 09-SEP-1998
ACCESSION	mgxb0021Dl19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0021Dl19r, DNA sequence.
VERSION	AQ163004
KEYWORDS	AQ163004.1 GI:3559405
SOURCE	GSS:
ORGANISM	Magnaporthe grisea. Magnaporthe grisea Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes Incertae sedis; Magnaporthaceae; Magnaporthe. 1 (bases 1 to 753)
REFERENCE	Yu, Y., Zhu, H., Boyd, C.A., Caudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A. A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome Unpublished (1998) Contact: Dean RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson University, Clemson, SC 29634 Tel: 864 656 5737 Fax: 864 656 4293 Email: rdean@clemson.edu Seq primer: GGAAACAGCTATGACCATG Class: BAC ends High quality sequence stop: 448.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

09-SEP-1998
 GSS
 DNA
 099 DP
 maxb0010f06r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mxgb0010f06r, DNA sequence.
 A Q162040
 A Q162040.1 GI:3558441
 GSS.
 Magnaporthe grisea.
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes Incertae sedis; Magnaporthaceae; Magnaporthe.
 1 (bases 1 to 699)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 285.
 Location/Qualifiers
 1. 699
 FEATURES
 source

BASE COUNT	208 A	180 C	156 G	206 T	3 others
ORIGIN					
Query Match	5.8%; Score 80.8; DB 225; Length 753;				
Best Local Similarity	48.8%; Pred. No. 7.1e-13;				
Matches 274; Conservative	0; Mismatches 280; Indels .7; Gaps 2;				
QY	90	ttgtgacactgttgacgggtgttaccaaatgttcccgagaaattctcacctgtgggggtac	149		
Db	189	TTTGTATGCGCCAGCCGGCTTTCAATTAAAGCAGCAATCACTCATAATTGGGGCCA	248		
QY	150	ctactctccatcattctcttttggcgacagaactcgtctattcttcgcagacgttccagac--	207		
Db	249	GTA TGACCATTTTTTCGTTGCCACGTCAGCACCATCGGAATTAAGAATTCCTCTGCTACCTT	308		

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9,926 clones with an average insert size of 130 kb was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

BASE COUNT	196 a	219 c	188 q	207 t
------------	-------	-------	-------	-------

Query Match 5.3%; Score 74; DB 225; Length 810;
Best Local Similarity 58.0%; Pred. No. 7.8e-11;
Matches 131; Conservative 0; Mismatches 95; Indels 0; Caps 0;

1179	Qy	cagttactctgctcttgactgcttccatcgcgtctagagcttacgttgaatgatgca	1233
1180			
1181			
1182			
1183			
1184			
1185			
1186			
1187			
1188			
1189			
1190			
1191			
1192			
1193			
1194			
1195			
1196			
1197			
1198			
1199			
1200			
1201			
1202			
1203			
1204			
1205			
1206			
1207			
1208			
1209			
1210			
1211			
1212			
1213			
1214			
1215			
1216			
1217			
1218			
1219			
1220			
1221			
1222			
1223			
1224			
1225			
1226			
1227			
1228			
1229			
1230			
1231			
1232			
1233	Qy	atgtcaagctaaaaaggaaaccatgtgttagagtttggttaaagacagagatttccatt	1298
1234			
1235			
1236			
1237			
1238			
1239			
1240			
1241			
1242			
1243			
1244			
1245			
1246			
1247			
1248			
1249			
1250			
1251			
1252			
1253			
1254			
1255			
1256			
1257			
1258			
1259			
1260			
1261			
1262			
1263			
1264			
1265			
1266			
1267			
1268			
1269			
1270			
1271			
1272			
1273			
1274			
1275			
1276			
1277			
1278			
1279			
1280			
1281			
1282			
1283			
1284			
1285			
1286			
1287			
1288			
1289			
1290			
1291			
1292			
1293			
1294			
1295			
1296			
1297			
1298			
1299	Qy	gcaagtggtgtgctgttgacaagttgggtagatgtaagagagacgacttcgttgaagttt	1358
1300			
1301			
1302			
1303			
1304			
1305			
1306			
1307			
1308			
1309			
1310			
1311			

[illegible]

JOURNAL COMMENT

Genome
Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Inst
Clemson University
100 Jordan Hall, Clemson Univer
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAAACACTATGACCATG
Class: BAC ends
High quality sequence stop: 465

FEATURES	source
Location/Qualifiers	
1. .669	
/organism="Magnaporthe grisea"	
/strain="70-15"	
/db_xref="taxon:148305"	
/clone="mgxb0004k10r"	
/clone_lib="CUGI Rice Blast BAC Library"	
/tissue_type="Protoplasts"	
/lab_host="E. coli DH10b"	
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII"	
Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with	

RESULT	11				
LOCUS	AV064132	234 bp	mRNA	EST	24-JUN-1999
DEFINITION	AV064132 Mus musculus small intestine C57BL/6J adult Mus musculus cDNA clone 2010009L02, mRNA sequence.				
ACCESSION	AV064132				
VERSION	AV064132.1	GI:5183960			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 234)				
AUTHORS	Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K., Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Kajiya,M., Kadota,K., Kagawa,I., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owai,C., Sato,K., Shibata,Y., Shigenoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N., Tsunoda,Y., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.				
TITLE	RIKEN Mouse ESTs				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Chie Owa Genome Science Laboratory RIKEN 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-9145 Fax: 81-298-36-9098 Email: genome-res@rtc.riken.go.jp Thermotabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998)) Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998)) Please visit our web site (http://genome.rtc.riken.go.jp) for further details.				
FEATURES	Location/Qualifiers				
Source	1..234				
	/organism="Mus musculus"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="2010009L02"				
	/clone_lib="Mus musculus small intestine C57BL/6J adult"				
	/sex="male"				
	/tissue_type="small intestine"				
	/dev_stage="adult"				
BASE COUNT	55 a	45 c	72 g	62 t	
ORIGIN					
Query Match	2.9%; Score 41.2; DB 109; Length 234;				
Best Local Similarity	52.3%; Pred. NO. 0.3;				
Matches	91; Conservative	0; Mismatches	83; Indels	0; Gaps	0;
QY	ctgttgaaatcattgaagaactgcaagtgcaggctgaacggaccattgtagagttt 1213				
Dd					
Db	CTGGTGGGCCCCCTCCACACTTGAGCTGCCTCGCTGGTGGTGGCTTCACCTCACTG 102				
QY	ctagagcttaagttgaaatgatcatgtcgaagtgaacggaccattgtagagttt 1273				
Dd					
Db	CGCGGCCGCCAGTTGAGGAGAAGAAAGTGGAGGGAACAAGAGATCCGAGGAGTTTG 162				
QY	tggttaaagcacagattgttccattgccagcggtgtgctgttgtgacaagttgggta 1327				
Dd					
Db	AAGATGACATGGGTCTCGGTTTTTTTGTGACTAAACTGCTTTTGTAAAGTTGGCTA 216				
RESULT	12				
LOCUS	AV293643	247 bp	mRNA	EST	09-NOV-1999
DEFINITION	AV293643 Mus musculus house muscle P1 mRNA, mRNA sequence.				
ACCESSION	AV293643				
VERSION	AV293643.1	GI:6307674			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 247)				
AUTHORS	Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ichii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kajiya,M., Kawaj,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owai,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigenoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y., Watanabe,S., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshioka,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.				
TITLE	RIKEN Mouse ESTs (Konno,H., et al. 1999)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Yoshihide Hayashizaki Genome Exploration Research Group, Life Science Tsukuba Center, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-res@rtc.riken.go.jp URL: http://genome.rtc.riken.go.jp/ Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsui,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y. Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998) Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site (http://genome.rtc.riken.go.jp) for further details.				
FEATURES	Location/Qualifiers				
Source	1..247				
	/organism="Mus musculus"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="5430440Cl6"				
	/clone_lib="RIKEN full-length enriched, 6 days neonate head"				
	/sex="mixed"				
	/tissue_type="head"				
	/dev_stage="6 days neonate"				
	/lab_host="DH10B"				
	/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'] GAGAGAGAAGGATCCAGAGCTCTTTTCTTTTCTTTTCTTTTCTTTN 3']; cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of				

100

us-09-488-265-28.rst

Sat Oct 27 15:25:46 2001

Result No.	Score	Query Match	Length	DB	ID	Description
1	1360.8	96.9	1426	9	AX021809	Sequence
2	798.8	56.9	1350	56	AF295325	Synthetic
3	635	45.2	1404	9	AF19452	phytase cDN
4	635	45.2	1404	10	I13430	Sequence
5	635	45.2	1404	10	I33881	Sequence
6	631	44.9	2000	13	ANPHYAG	A.niger phy
7	631	44.9	2665	13	ANSPHYTASE	Aspergillus
8	631	44.9	6756	9	AF19451	phytase gen

[illegible]

QY	1141	ttgtctactactctgttgaatactattgaagaactgcagcggttactctctcttggaact	1200
Db	1152	TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTCTTGGACT	1211
QY	1201	gtccattcgcctagacttacgttgaaatgatgaatgataagctgaaaggaaacca	1260
Db	1212	GTTCATTCGGTGCTAGACCTTACGGTTGAATGTGAATGATGCAATGTCAAGCTGAAAAGGAACCA	1271
QY	1261	ttggttagagtttbggttaacacacagagtgttccattgcacggttgctgctgttaacaag	1320
Db	1272	TTGGTTAGAGTTTGTGTTAAACACACAGAGTGTTCATTGCACGGTTGCTGCTTGACAAG	1331
QY	1321	ttgggttagatgaacagagacgaactcgttgaaagtttgcttagatcctgtagatcgtggt	1380
Db	1332	TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTCCTCGCTAGATCTGGTGT	1391
QY	1381	aactgggctgaagtcttcgcttaa	1404
Db	1392	AACTGGGCTGAATGTTTCGCTTAA	1415
RESULT	2		
AF295325	AF295325	1350 bp mRNA	SYN 18-OCT-2000
LOCUS	Synthetic construct phytase mRNA, complete cds.		
DEFINITION	AF295325		
ACCESSION	AF295325.1	GI:10732782	
VERSION			
KEYWORDS			
SOURCE	Synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequence.		
AUTHORS	1 (bases 1 to 1350)		
TITLE	Yang, L., Chen, Z., Bei, J., Liao, L. and Wang, X.		
JOURNAL	Synthetic sequence of phytase gene for expression in Pichia		
REFERENCE	pastoris		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 1350)		
JOURNAL	Chen, Z.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (11-AUG-2000) Animal Science Institute, Guangdong Academy		
TITLE	of Agricultural Sciences, Guangzhou, Guangdong 510640, China		
JOURNAL	3 (bases 1 to 1350)		
REFERENCE	Yang, L., Bei, J., Liao, L. and Wang, X.		
AUTHORS	Direct Submission		
TITLE	Submitted (11-AUG-2000) Biopharmaceutical Center, Zhongshan		
JOURNAL	University, Guangzhou, Guangdong 510275, China		
FEATURES	Location/Qualifiers		
source	1..1350		
	/organism="synthetic construct"		
	/db_xref="taxon:32830"		
	/note="Aspergillus niger sequence modified for expression		
	in Pichia pastoris"		
CDS	1..1350		
	/codon_start=1		
	/transl_table=11		
	/product="phytase"		
	/protein_id="AAG22536.1"		
	/db_xref="GI:10732783"		
	/translation="MLAVPASRNQSCDVTDOGYQCFSETSLHWQYAPFFSLANESV		
	ISPEVPAGDCTFFVAGLRSRGARYPTDSKGYKSALIEEIQONATTFGKYAFLKTYN		
	YSLGADDTPTFGEQLVNSGKIFYQYVESLFTNRNIVFIRSSGSRVIAAGKKFIEGQ		
	STKLDPRAQPCQSSPKIDVVISSESSNLTDPCTCFEEDSELADTVEANFTATFY		
	PSIQRLEDNLGVTLTFTVYTLMDGMSDTISTSTVDDTKLSPKCDLTFHDEWNLNY		
	STLVADESHDNGITSILFALGLNGTKPLSTTVENITOTDGFSSAWTVPFASRLYVE		
	MMQCAQEPLVRLVINDRVVPLHGCPVDALGRCTRDTSFVRGLSFARSGGWAECFA"		
BASE COUNT	319 a 354 c 282 g 395 t		
ORIGIN			
Query Match	56.9%	Score 798.8; DB 56; Length 1350;	
Best Local Similarity	75.0%;	Pred. No. 4.4e-206;	
Matches 998; Conservative	0; Mismatches 332;	Indels 0; Gaps 0;	

QY 1155 tggtaattatgaagaactgacggttactctgtctcttctgactgttccattcgtgc 1214
 Db 1101 CGTTGAGAAATATCACCCAGACAGATGATCTCTCTGTTGGACTGTCCATTGCTTC 1160
 QY 1215 tagagcttactgtgaatgatcaatgtcaagctgaaaggaacaccattggttagagattt 1274
 Db 1161 TAGATTGACGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220
 QY 1275 gtttaacacagagattgttccattgacggttctgtctgttgaagaagttggttagatgtaa 1334
 Db 1221 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1280
 QY 1335 gagagcagcattctgtgaaggttcttctgtctgtctgtctgtctgtctgtctgtctgt 1394
 Db 1281 CAGAGATCTTTGTTAGAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1340
 QY 1395 tttegcctaa 1404
 Db 1341 TTTTGCCTAA 1350

RESULT 3
 A19452 1404 bp mRNA PAT 10-JUN-1994
 LOCUS
 DEFINITION phytase cDNA fragment.
 ACCESSION A19452
 VERSION A19452.1 GI:583195
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 REFERENCE
 1 (bases 1 to 1404)
 AUTHORS van Gorcom, R.F.M., van Hartingsveldt, W., van Paridon, P.A.,
 Veenstra, A.E., Luitken, R.G.M. and Selten, G.C.M.
 TITLE Cloning and expression of microbial phytase
 JOURNAL Patent: EP 0420358-A 41 03-APR-1991;
 GIST-BROCADES N.V.

FEATURES
 source
 1. .1404
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 gene
 1. .1404
 /gene="Phytase gene"
 CDS
 1. .1404
 /gene="Phytase gene"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA01464.1"
 /db_xref="GI:583196"
 /translation="MGYSAVLLPLYLLSGVTSGLAVPASRNQSCDVTDOGYQCFSET
 SHLWQVAPFSLANESVISPVEAGCRVTFQVLSRHGARYPTDYSKYSALIEPI
 QONATFDKTAFLKTNYSILGADLTFRGEQLVNSGIKFYORFSLRNIVPFTRS
 SGSRVITASGKKFIEGFQSTKLKDPRAQPGQSSPKIDVISEASSNNITLDPGCTVF
 EDSLEADTVANFTAFVPSRQRLNDLSGVTLTDTEVYLMDCMFSFRTSTVDFT
 LKSPCDLTHDEWINDYQLSKYKCHGAGNPLGPTGVGVANELLARTLSPVHD
 DTSNNITLSDSPALINDYLADEFSHDNGLIISLFAIGLYNGTKPLSTTTVENITQT
 DGFSSAWTPPASRLYVEMMQCAEQEPLVRVLVNDVRVPLHGCPVDALGRCTRDSEV
 RGLSFARSGGDWACFA"
 BASE COUNT 293 a 436 c 344 g 331 t
 ORIGIN

Query Match 45.2%; Score 635; DB 9; Length 1404;
 Best Local Similarity 65.8%; Pred. No. 1,3e-161;
 Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;
 QY 1 atggcggtgtgtgctgactgtccattgacacattgttgcgttccacatcccggtacc 60
 Db 1 ATGGCGGTCTGCTGCTTCTACTCTCTTTGTTATCTCTCTGTTGGAGTCACTCCGCGACTG 60
 QY 61 gcttgggttctgtgtgtaattctactctgttgcacactgttgcaggtgtggttaccatgt 120
 Db 61 GCAGTCCCCCGCTCGAGAAATCAATCCAGTTGCGATACGGTCTGATCAGGGGTATCAATGC 120

QY 121 ttcccagaaatttctcacttctgtgggttaactctcactctcactctcactctcactctcactctc 180
 Db 121 TTCTCCGAGAGATTTGCGATCTTTGGGTCATACGACGACGATCTCTCTCTCTCTCTCTCTCTCTCT 180
 QY 181 tctgtatttctcagagcttccagagcagctgttagagttacttcttctcgttcaagttttgtct 240
 Db 181 TCGGTTCATCTCCCTGAGGTGCGCGCGGATGCAGAGTCACTTTTCGCTCAGGTCCTCTCC 240
 QY 241 agacacggtgctagataccacaaacttctctgctgctctaaaggttactctctgtttgattgaa 300
 Db 241 CGTCATGAGCGCGGTATCCGACGACCTCCAAAGGCGCAAGAAATACTCCGCTCTCATTTGAG 300
 QY 301 gctattcaaaagacgctactgtttcaaggttaaglacgcttcttcttgaagacttacaac 360
 Db 301 GAGATCCAGCAGACGCGACCCCTTTGACGGAAATATGCTTCTCTGAGACATACAAC 360
 QY 361 tacactttgggtgctgacgactgtactctcattcgttgaacacacaaatgggttaactcgtt 420
 Db 361 PACAGCTTGGGTGCAGATGACCTGACTCCCTTCGGAACACAGAGAGTAGTCAACTCCGCGC 420
 QY 421 attaagttctacagaagatacaaggtttggcttagaagattgttccattcatttagagct 480
 Db 421 ATCAAGTTTCTACACGCGGTACGAATCGCTCAAGGAAACATCGTTCCATTCTATCCGATCC 480
 QY 481 tctgttctgacagagttattgtcttctgctgctgaaagttcattgaaggtttccaatctgct 540
 Db 481 TCTGGCTCCAGCGCGTGTGCTCCGCGAAGAAATTCATCGAGGGGCTTCAGAGCACC 540
 QY 541 aagttggtgacccaggttctcaac 600
 Db 541 AAGCTGAAGGATCTCGTCCG 600
 QY 601 ccagaagatccggttacaac 660
 Db 601 TCGGAGCGCAGCTCATCCAAACACACTCTCGACCCAGGACCTGCGACATGTCTTCGAAGAC 660
 QY 661 tctgaattaggtgacgactgtgaagcttaactctactctgttctgctcagcttacttaga 720
 Db 661 AGCGAATTGGCGGATACCGCTCGAAGCCAAATTCACGCGCGCACCGTCTGCTCCCTCCATTCGT 720
 QY 721 gctagattgaagcagctactgacaggtgttacttgaactgacgacgacgacgacgacgacgac 780
 Db 721 CAACGCTGGGAGAACGACCTGTCCGGGTGACTCTCACAGACACACAGAGTACCTACCTC 780
 QY 781 atggacatgtgtccattgacacactgtcgttagaacttctgacgctactgaaattgtctoca 840
 Db 781 ATGGACATGTGCTCTTCGACACCACTCTCCACGACCGCTCGACACCAAGGTGTCCGCC 840
 QY 841 tctgtgttctcactcagcagcgaatggatcccaatacgaactacttgcgaagcttgggt 900
 Db 841 TTCTGTGACCTGTTCACCCATGACGAATGGATCAACTACGACTACCTCCGCTCTCTTGAAA 900
 QY 901 aagtactcaggttacgctgctgtgtaacccattgggttccagctcaaggtgtgtgttctcgt 960
 Db 901 AAGTATTACCGCCATGGTGGAGTAACCCGCTCGCGCCGCGACCCAGGCGTTCGGCTACGCT 960
 QY 961 aacgaattgattgctagattgactcactctcagtttcaagacacacacttctactaacaac 1020
 Db 961 AACGAGCTATCGCCGCTGTGACCCACTCTGCTCCGCTGTCCAGATGACACACCAAGTTCACAC 1020
 QY 1021 actttgactctaaaccagctacttctccacttgaacgctacttctgacgctactctct 1080
 Db 1021 ACTTTGAGACTCGAGCGCGGTACTCTTTCGCTCAACTCTACTCTACTCTACGGGGACTTTTCG 1080
 QY 1081 cagcacaactactgatctattttctctgcttgggtttgttacaacggttaccagca 1140
 Db 1081 CATGACAAACGGCACTCATCTCACTCTCTTCTTTAGTCTGTACAAACGCGACTAAAGCG 1140
 QY 1141 ttgtctactacttctgttgaatctattgaagaactgaacggttactctgcttctgagact 1200
 Db 1141 CTATCTACCCAGCGGTGGAGAAATATACCCAGACAGATGGATTCGCTCTGCTCTGTGGACG 1200

KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 BASE COUNT 293 a 436 c 344 g 331 t
 ORIGIN
 Query Match 45.2%; Score 635; DB 10; Length 1404;
 Best Local Similarity 65.8%; Pred. No. 1.3e-161;
 Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;
 QY 1 atggcggtgtctgtctactgtccacttgcacactgttgcggttccacatcccggtacc 60
 DB 1 ATGGGCGTCTCTGCTGTCTACTTCTTTGGTATCTCTGCTGTGGAGTCACTCCGGACTG 60
 QY 61 qcttgggtctcgtgtgaattctcactctctgtgacactgttgaacggttgaacaaatgt 120
 DB 61 GCAGTCCCGGCTCGAGAAATCAATCCAGTTCGATACGTCATCAGGGGTATCAATGC 120
 QY 121 ttccagaaattctcacttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 DB 121 TTTCTCCGAGACTTCGATCTCTTGGGTCAATACGCCACCTCTTCTCTGTGCAACGAA 180
 QY 181 tctgtatttctcagacgttccagacactgtagagttacttctgttcaagtgttctgt 240
 DB 181 TCGGTCAATCTCCCTGAGGTGCGCGGATGCAGAGTCACTTTCGTCAGTCTCTCTCC 240
 QY 241 agacacggtgtagataccacaaattcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 DB 241 CGTCATGAGCGGGTATCCGACCGACTTCCAGGCGAAGAAATCTCCGCTCTCATTTAG 300
 QY 301 gctattcaaaagacgtactgttctcaggtgaaggtgaaggtgtgtgtgtgtgtgtgtgt 360
 DB 301 GAGATCCAGACAGCGGACCCCTTTCAGCGAAATATGCCCTTCTGAGACATACAC 360
 QY 361 tacacttgggtgtgacgtactgtacttccacttgcgtgtgtgtgtgtgtgtgtgtgtgt 420
 DB 361 TACAGTTGGGTGAGATGACCTGACTCTCCCTTCGGAGAACAGGAGTACTCAACTCCGC 420
 QY 421 attaatgtctacagaatatacagaagcttgggtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
 DB 421 ATCAAGTTCTACAGCGGTAGAAATCGCTCACAAGGAACATCGTTCATTCATPCGATCC 480
 QY 481 tctgttctgacagagttattgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
 DB 481 TCTGCTCCAGCGGTGATCGCTTCGGGCAAGAAATTCATCAGGAGTTCAGAGACCC 540
 QY 541 aagttgtgtgacccaggttctcaaccacacacacacacacacacacacacacacacacac 600
 DB 541 AAGCTGAAGGATCCTCGTCCGACCGCGGCAACATCGTCGCCCAAGATCGAGTGGTCAAT 600
 QY 601 ccagaagatccggtttac 660
 DB 601 TCCGAGGCCAGCTCATCCAAACACACTCTGACCCAGCGACCTGCTGCTGCTGCTGCTGCT 660
 QY 661 tctgaattagtgacgacgttgaagtaacttcaactgttctgtgtgtgtgtgtgtgtgtgtgt 720
 DB 661 AGCGAATTGGCGGATACCGTCGAGGCCAATTTACCCGCCACGTTCTGCTGCCCTCCATTCGT 720
 QY 721 gctagattggaagctgacttccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
 DB 721 CAACGCTGGAGAACGACTCTGCGGGTGTACTCTCACAGACACAGAAAGTACGCTACCTC 780
 QY 781 atgacatgttccattcagacactgtcactgaagacttctgacgtcactgaattgtctcca 840

```

exon      166..253
/feature="phyA"
/number=1
gene      166..1871
/feature="phyA"
prim_transcript 166..1871
/partial
/feature="phyA"
/function="primary transcript of phyA gene"
join(210..253,356..365)
/feature="phyA"
CDS      join(210..253,356..1715)
/feature="phyA"
/codon_start=1
/protein_id="CAA78904.1"
/db_xref="GI:2393"
/translation="MGVSALLPLYLISGTVSLGAVPASRNQSSCDTVDQGYCFSET
SHLWQYAPFFSLANESVISPEVACGRVTFACVLSRHGARYPTDSKKGKYSALIEEI
QONATFDGKYAFLEKTYNSYSLGADDLTPFGEQELVNSGIFKFORYESLTRNVPFIRS
SGSRVIAAGKXKIEGFQSTKLKDPRAQPGSQSPKIDVYSEASSNNTLDPTCTVF
EDSELADTVEANFTAFVPSIRORLENDLSGVLTDTEVYILMDWCSDDTISTSVTD
KLSPFCDLTHDEWINDYLOSLLKYYGHGACNPLGPTQGVGYANELLARUHPHVD
DTSSNHTLDSPPATPLNSTLYADESHDNGIISILFALGVLNGTKPLSTTVENTQT
DGFSSAWTPFASRLYVEMMQAQOEPLRVRLVNDVRVPLHGPCVDALGRCTRDSFV
RGLSFARSGDWAECFA"
intron    254..355
/feature="phyA"
/number=1
356..1878
/number=2
mat_peptide 369..1712
/feature="phyA"
stem_loop 1732..1760
/feature="phyA"

BASE COUNT 463 a 603 c 456 g 478 t
ORIGIN

Query Match 44.9%; Score 631; DB 13; Length 2000;
Best Local Similarity 65.9%; Pred. No. 1.7e-160;
Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;

QY 13 gtcgtgctactgtccattgccaactgtgttcggttccacatccggtacacgcttggtgct 72
DB 324 GTGGGACTACTGATCGCTGACTATCTGTGCAGAGTCACCTCCGGACTGGCAGCTCCCGCC 383

QY 73 cgtgtaattctcactctgtgacactgttgacggttgacggtgttaccgaatgttccagaatt 132
DB 384 TCAGAGAAATCAATCCAGTTGGCATACGGTGCATAGGGGTATCAATGCTTCTCCGAGACT 443

QY 133 tctcaacttgggggtacctactctccacttctcttcttggcagacgaatctgctatttct 192
DB 444 TCGCAATCTTTGGGGTCAATPACACACCGTCTTCTCTCTGGCAACCAATCGGTCTATCTCC 503

QY 193 ccagacgttcccgacgactgtagacttcttcttcttcttcttcttcttcttcttcttctt 252
DB 504 CCTGAGGTCCCGCGGATGCAGATGACTTCTTCTGCTCAGGTCTCTCCCGCTCATGGAGCG 563

QY 253 agatacccaacttcttctgctgtaagcttactctgcttcttcttcttcttcttcttcttct 312
DB 564 CGGTATCCGACCGACTCCAAGGGCAAGAAATAC1CCGCTCTCATTTGAGGAGATCCAGCAG 623

QY 313 aacgctactgtttcaagggttaagtacgcttcttcttcttcttcttcttcttcttcttctt 372
DB 624 AACGGACCACTTTGACGGGAAATATGCTTCTCTGAAGACATACAACTACAGCTTGGT 683

QY 373 gctgacgacttactccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 432
DB 684 GCAGATGACTGACTCTCCGAGACAGGAGTAGTCAACTCCGGCATCAAGTTCTAC 743

QY 433 agaagatacaagggttggctgagaaagattgttccattcattagagcttctggttctgac 492

```

RESULT 7
 ASNPHYTASE
 LOCUS
 DEFINITION
 Aspergillus niger myo-inistol hexaphosphate phosphohydrolyase gene,
 complete cds.
 M94550

PLN

27-APR-1993

DNA

2665 bp

ASNPHYTASE

Aspergillus niger

myo-inistol

hexaphosphate

phosphohydrolyase gene,

complete cds.

M94550

Db 744 CAGCGGTACGAATCGCTCACAGGAACATCGTTCCATTTCATCCGATCCTCTGCTCCAGC 803
 QY 493 agagttattgtcttctgctgaaagtattcaaggtttccaaactgctgaagtgtgctgac 552
 Db 804 CGCGTGTATCGCTCCGGCAAGAAATTCATCGAGGGGCTCCAGAGCACCAGCTGAAGAT 863
 QY 553 ccaggttctcaaccacacccaagcttctccagttatttaacgtgtgattcattccagaagatcc 612
 Db 864 CTTCTGTGCGCCAGCCGCCCAATCGTCGCCCAAGATCGAGCTGCTATTTCAGGCCAGC 923
 QY 613 ggttacacacacacttggaccacggttactgtactgttcttgcgaagactctgaatagt 672
 Db 924 TCATCAACACACACTCTCGACCCAGCAGCACTGCTACTGTCTTCGAAGACAGCAATGGCC 983
 QY 673 gacgagcttgaagctaaacttactgcttcttctgctccagctatttagagctagattgaa 732
 Db 984 GATACCGTCGAAGCAATTTACACGCCAGCTTCGTCCTCCATTCGTCACAGCTCTGGAG 1043
 QY 733 gctgacttgcaggtgttacttgcagacgaagacgttcttactgtatgacatgtgt 792
 Db 1044 AACGACCTGTCCGGTGTGACTCTCAGACACAGAAAGTGAAGTACTTCTCATGACATGTC 1103
 QY 793 ccattcgacactgtcgctagaaacttctgacgtactgaatgtctccattctgtgcttgg 852
 Db 1104 TCCTTCGACACCATCTCCACAGCACCGCTGACACCAAGCTGTCCCTCTCTGTGACCTG 1163
 QY 853 ttcactcacgaaatggtatccaatcacgactacttgcgaagcttggtaagtactacggt 912
 Db 1164 TTCACCATGAGAAATGATGATCACTACGACTACCTCCAGTCTTGAANAAGTATTACGGC 1223
 QY 913 tacggtgctgtaaccatttgggtccagctcaaggtgtgttcttgcgaagactgattgatt 972
 Db 1224 CATGTTGACAGTAAACCCGCTCGGCCGACCCAGGGCGCTCGGCTACGTTACGAGTCTATC 1283
 QY 973 gctagattgactcactctccagttccaagacacacttctactaaccacactttgactct 1032
 Db 1284 GCCGCTGACCCACTCGCCTGTCCAGATGACACAGTTCCAAACCACTTTTGGACTCG 1343
 QY 1033 aaccagctacttctccattgacgctacttctgacgtgacttctctcagcaacaact 1092
 Db 1344 AGCCCGGCTACTCTTCCGCTCAACTACTCTCTACGCGGACTTTTCGCATGACAACGGC 1403
 QY 1093 atgatatctatttctcgtcttgggtttgtacacaggttaccacagcactgtctactact 1152
 Db 1404 ATCATCTCCATCTCTTGTCTTAGGTCTGTACAAACGGCACTAAGCCGCTATCTACCACG 1463
 QY 1153 tctgttgaatctattgaagaaactgacggttactctcttcttggactgttctcattcgt 1212
 Db 1464 ACCGTGGAGAAATATCACCCAGACAGATGGATCTCTGCTCTGCTGGAGGTTCCGTTGCT 1523
 QY 1213 gctagagcttactgtgaatgatgcaatgtcaagctgaaaggaacacatttggttagatt 1272
 Db 1524 TCGGTTGTGACGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
 QY 1273 ttggttaacgacagagttgttccattgacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1332
 Db 1584 TTGTTTAATGATCGGTGTGTCGCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1643
 QY 1333 aagagagacacttctgtgaaggttcttctcgttagatctgtgtgtgtgtgtgtgtgtgtgt 1392
 Db 1644 ACCCGGATAGCTTGTGAGGGGTTGAGCTTGTGCTTAGATCTGCGGGTGTGATTTGGCGGAG 1703
 QY 1393 tgttctgctta 1403
 Db 1704 TGTTTGTCTTA 1714

[illegible]

Query Match	44.9%;	Score 631;	DB 9;	Length 6756;
Best Local Similarity	65.9%;	Pred. No. 1.8e-160;		
Matches 916;	Conservative 0;	Mismatches 475;	Indels 0;	Gaps 0;
QY	13	gtcgtgctactgtcccaattgccacctgtgttcgattccacatccggtaccgccttgggtcct	72	
Db	324	GTGGGACTACTGATCGCTGACTATCTGTCCAGAGTCACCTCCGGAGTCGCGAGTCCCGGCC	383	
QY	73	cgtggttaattctcaactctgtgtgacacgtgtgaocgttggtttaccaatgtttccccagaatt	132	
Db	384	TCGAGAAATCAATCCAGTTTGGCATAGGTCGATCAGGGGTATCAATGCTTCCGAGACT	443	
QY	133	tctcaactgtgggggtacctactctccactctctctcttcttggcgacgacgaatctcgtattct	192	
Db	444	TCGCATCTTTGGGTCAATFACGACCGTCTTCTCTCTCGCAACGAATCGGTCACTCC	503	
QY	193	ccagacggttccagacgactgtagagttacttctcgttccaagttttgtctctagacacggtgct	252	

Db 1584 TTGCTTAATGATCGCTTGTCCCGCTGCATGGGTGTCCGGTTGATGCTTTGGGAGATGT 1643
 Qy 1333 aagagagacacttcggtgaaggtttgttcctcgtatagatcgtgtgtaactgggtgaa 1392
 Db 1644 ACCCGGGATAGCTTTGTGAGGGGTTGAGCTTGTCTAGATCTGGGGGTGATTTGGCGGAG 1703
 Qy 1393 tgttcgctta 1403
 Db 1704 TGTTCGCTTA 1714

RESULT 9
 LOCUS I13429 6756 bp DNA PAT 26-JUL-1995
 DEFINITION Sequence 31 from patent US 5436156.
 ACCESSION I13429
 VERSION I13429.1 GI:910770
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE
 1 (bases 1 to 6756)
 AUTHORS Van Gorm, R.F.M., Van Hartingsveldt, W., Van Paridon, P.A.,
 Veenstra, A.E., Luiten, R.G.M., and Sellen, G.C.M.
 TITLE Cloning and expression of phycase from aspergillus
 JOURNAL Patent: US 5436156-A 31 25-JUL-1995;
 FEATURES
 Location/Qualifiers
 1..6756
 /organism="unknown"

BASE COUNT 1615 a 1714 c 1712 g 1715 t
 ORIGIN

Query Match 44.9%; Score 631; DB 10; Length 6756;
 Best Local Similarity 65.9%; Pred. No. 1.8e-160;
 Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;

Qy 13 gtcgtgtactgtccatgccaactgtgttcgtccacatccggtacccgcttgggtct 72
 Db 324 GTGGGACTACTCGCTGACTATCTGTGACAGTCACTCCGGACTGGAGTCCCGCC 383
 Qy 73 cgtggttaattctcactctgtgacactgtgacgtgtgttaccatgtttccacgaatt 132
 Db 384 TCGAGAAATCAATCCAGTTGGCATACGGTTCGATCAGGGGTATCAATGCTTCTCCGAGACT 443
 Qy 133 tctcaactgtgggttacactctccatctctcactctctctgttggcagacgaatctgtattct 192
 Db 444 TCGCATCTTTGGGGTCAATACACCGCTTCTCTCTGGCAACGAATPCGCTCATCTCC 503
 Qy 193 ccagacgttccagacgactgttagagttacttctcgttcaagttttgtctagacacggtgt 252
 Db 504 CFTGAGGTGCGCGCGGATGACAGTCACTTTGCTCAGGTCTCTCCCGTCATGGAGCG 563
 Qy 253 agatacccaactctctcgtcgttaaggtctactctgtgttattgattgaagctattcaaaag 312
 Db 564 CGGTATCCGACCGACTCCCAAGGCAAGAAATACCTCGCTCTCAATTTGAGGAGATCCAGCAG 623
 Qy 313 aacgtactgtttcaaggttaagtagctgttcttgaagacttaacactacactttgggt 372
 Db 624 AACGCGACCACTTTGACGGAAATATGCTTCTTCCCTGAAGACATACACTTACAGCTTGGGT 683
 Qy 373 gctgacgacttgactccatcgtgtgaaacacaaatggttaactcgtgttattgaagttctac 432
 Db 684 GCAGATGACTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 743
 Qy 433 agaagatacaaggtcttggctagaagattgttccattcattagactcttctgtctgac 492
 Db 744 CAGCGGTACGAATCGCTCAACGAAGACATCGTTCCATTCATCGATCTCTGGCTCCAGC 803
 Qy 493 agagtattctctcgtgaaagttcaatgaaagtttccaatctgcttaagttggctgac 552
 Db 804 CCGTGTATCGCTCCGCAAGAAATTCATCGAGGGCTCCAGAGCACCACCAAGCTGAAGGAT 863

Qy 553 ccaggttctcaaccacacacaaagctttctccagttattaaagtgatcattccagaagatcc 612
 Db 864 CCTCGTCCACACCGCGCAATPCGTGCGCCCAAGATPCGACGGTGGTCAITTCGAGCGCAGC 923
 Qy 613 ggttacaacaacactttgaccacggttactgtactgtttctcgaagactctgaattaggt 672
 Db 924 TCATCCACAACACTCTGACCCAGCAGCTGCACTGTCTTCGAAGACAGCAATTTGCC 983
 Qy 673 gacgacgttgaagctaaacttoactcgtttgttgcgtccagctatttagagctagattgaa 732
 Db 984 GATACCGTCGAAGCCAATTTACCCGCCACGTTGTCGCCCTCCATTCGTCACGCTGGAG 1043
 Qy 733 gctgacttcacaggttacttttgactgaagacggtttgttacttgatgacatgtgt 792
 Db 1044 AACGACCTGTCGGGTGACTCTCACACACAGAGTACCTACTCTATGGACATGTC 1103
 Qy 793 ccattcgacactgtcgttagaactttgacgctactgaattgtctccattctgtgctttg 852
 Db 1104 TCCTTCGACACCATCTCCACCAGCAGCCGTCGACACAAAGCTGTCCCTTCTGTGACCTG 1163
 Qy 853 ttoactcaagacgaatggtatccaatacagactacttgcaagcttgggtaagtactacgt 912
 Db 1164 TTCACCCATGAGATGATGATCACTACGACTACCTACCTCCAGTCTTGAAGAAATTTACGGC 1223
 Qy 913 tacggtgtggttaacccattgggtccagctcaaggtgtgtgttttcgcgttaacgaattgatt 972
 Db 1224 CATGGTGCAGGTAAACCGCTCGCGCGACCCAGGGCTCGGCTAACGAGCTCATC 1283
 Qy 973 getagattgactcactctccagttcaagacacacactcttactaaacacactttgactct 1032
 Db 1284 GCCCGTCTGACCCACTCGCTCTCCACGATGACACAGTTCACACCCACACTTTTGACTCG 1343
 Qy 1033 aacccgacttctccattgaacgctactttgtacgctgacttctctcaacacacact 1092
 Db 1344 AGCCGGCTACTCTTCGGCTCAACTCTACTCTACGCGGACTTTTCGCATGACAACGGC 1403
 Qy 1093 atgatatctatttctcgttgggtttgtacacacggttaccacacacactgtctactact 1152
 Db 1404 ATCATCTCCATCTCTTGTGTAGTCTGTACAAACGGCACTAAAGCCGTATCTACACAG 1463
 Qy 1153 tctgtgaactctattgaagaaactgacggttactctgtcttctgtgactgttccactcgt 1212
 Db 1464 ACCGTGAGAAATATCACCAACAGATGATTCGCTCTGCTGACGCTTCGTTTGTCT 1523
 Qy 1213 gctagagcttaccgttgaatgatgcaatgcaagctgcaagctgaaagacacacatttggtagatt 1272
 Db 1524 TCGGTTTGTGACGTCGAGATGATGCACTGTACGGCGAGCAGGAGCGCTGTCTCCGTGTC 1583
 Qy 1273 ttggttaacacagagattgttccattgacggttgcgtgtgacaaagttgggttagattg 1332
 Db 1584 TTGTTTAATATCGCGTGTCCCGCTGCATGGGTGTCCCGTTGATGCTTTGGGGAGATGT 1643
 Qy 1333 aagagacgacttcgttgaaggtttgtcttctcgttagatctggtgtaactgggtgaa 1392
 Db 1644 ACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGCGGAG 1703
 Qy 1393 tgttcgctta 1403
 Db 1704 TGTTCGCTTA 1714

RESULT 10

AF218813
 LOCUS Aspergillus niger myo-inositol hexaphosphate phosphohydrolyase
 DEFINITION precursor, gene, complete cds.
 ACCESSION AF218813
 VERSION AF218813.1 GI:6694940
 KEYWORDS Aspergillus niger.
 ORGANISM Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE

- 1 (bases 1 to 1528)
HONGNING, W., QI, W. and JING, X.
PCR, cloning and characterization of the phytase (phyA) gene of
Aspergillus niger (China Strain)
Unpublished
2 (bases 1 to 1528)
HONGNING, W., QI, W. and JING, X.
Direct Submission
Submitted (24-DEC-1999) Aniaml Sci. & Tech. College, Sichuan
Agricultural University, Ya'an City, Sichuan Province 625014,
P.R.China

FEATURES

source
1. 1528
/organism="Aspergillus niger"
/db_xref="taxon:5061"
/country="China"
join(14..57,160..172)
join(14..57,160..172)
/product="myo-inositol hexaphosphate phosphohydrolyase
precursor"
join(14..57,160..1519)
/EC_number="3.1.3.8"
/function="catalyzes the hydrolysis of phytic acid"
/note="phytase"
/codon_start=1
/product="myo-inositol hexaphosphate phosphohydrolyase
precursor"
/protein_id="AAF25481.1"
/db_xref="GI:6694941"
/translation="MGVSAVLLPLVLLSGVSGGLAVPASRNOSTCDVVDGQYCFSET
SHLWQYAPFFSLANESALSDVPAGCRVTFQVLSRHGARYPTDSKGRYSALIEEI
QNAATFDKAYFLKTYNSLGGADLTPFGEQLVNSGKIFQYVESLIRNIIPFIRS
SGSRVIAAGKRIEFGOSTKLDPRAQPSQKIDVVISEASSNMLDPGCTVFP
EDSELADAEANFTATFVPTIRORLENDLSGVSLTDTEVYLMQMSFDTSTSTVDI
KLSPCDLFTHDENYNDYLOSLLKYKHGAGNELGPTQGVGYANELLIRATHSPVHD
DTSSNHTLDSASFTPLNSTLYHAFSDHNGLIISILFALGLYNGTKPLSTTVQNIOT
DGFSANWYVPSRLYIVEMOCQAEPLVRLVNDRVVPLHGGPVDALGRCTRDSFV
KGLSPARSGGDWAECEFA"

BASE COUNT 326 a 475 c 364 g 363 t
ORIGIN

Query Match 44.8%; Score 629.4; DB 13; Length 1528;
Best Local Similarity 65.8%; Pred. No. 4.5e-160;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

QY 13 gtgtgtactgtccattgccacctgtgtccattccacatccggtaccgcttgggtct 72
Db 128 GTGGGACTACTGATCGTGACCATCTGTCCAGAGTCACCTCCGGACTGGCAGTCCCGGCC 187
QY 73 cgtgtaattcactctgtgacactgtgacggtgtgacggtgttaccagtgttcccgagaaatt 132
Db 188 TCAGAAATCAATACACATTCGATACGGTTCGATACGGGTATCAATGCTTCTCCGAGACT 247
QY 133 tctcactgtgggttacactctctcactctctcactctcttcttggcagacgaatctgtattct 192
Db 248 TGGCATCTTTGGGTCATACGGCCGCTCTTTCTCTGGCAACGAATCGGCATCTCT 307
QY 193 ccagacgttccagacgacttagagttacttcttccagtgttccagtgttccagtgttccagt 252
Db 308 COTGATGTGCCCGCGGATCGGAGTCACTTTCGCTCAGGTCCTCTCCGCTCATGGAGCA 367
QY 253 agatacccaactctctctgctctaaagcttacttcttccagtgttccagtgttccagtgttccagt 312
Db 368 CGGTATCCGACCGACTCCAAAGGCAAGAAATPACTCCGCTCTCATTTGAGGAGATCCAGCAG 427
QY 313 aacgtactgtcttcaaggggtaagtaacgcttcttcttgaagacttaacactacactttgggt 372
Db 428 AATGCGACCACTTTGATGGGAATATGCCCTTCTTGAAGACATACAACTACAGCCTGGGT 487
QY 373 gctgacgacttgactcattcgttgaaacacaaatggttaactctgtatttaagttctac 432
Db 488 GCAGATGACCTGACTCTCTTCGGAGAACAGGAGCTAGTCAACTCCGSCATCAAGTCTAT 547

QY 433 agaagatacaaggcttggcttagaagaattgttccattcattagagcttctgtgttctgac 492
Db 548 CAGCGATACGATCGTCTCACAAGAAACATCATTCATTCATTCCTCTCGCTCCAGC 607
QY 493 agagttattgtctgtctgctgaaagttcattgaaggtttcccaatctgctaaagtgtgctgac 552
Db 608 CCGCTGATCGCTCCGGCAAGAAATTCATCGAGGCTTCCACAGCAGCAACGCTGAAGAT 667
QY 553 ccaggttctcaaccacacacagcttctccagttattaaagtgatcattccagaagatcc 612
Db 668 CTTCTGCGCCAGCCCGCCCAATCGTCGCCAAGATCGAGTGGTGTATTCGAGGCCAGC 727
QY 613 ggttaacaacacactttggaccacggtactgttactgtcttctgagagactctgaattaggt 672
Db 728 TCATCCAAACACACTCTCGACCCAGGACCTGCCTGTCTTGAAGACAGGAAATGGCC 787
QY 673 gacgagttgaagctaaacttaacttaactgttctgtctcagctatttagagcttagattgaa 732
Db 788 GATCGCGTCGAAGCCAAATTCACCGCCACGTTCTGTCGCCACCATTCGTCACAGCTGGAG 847
QY 733 gctgacttgcaggtgttacttgaactgaagacggtgttcttactgtatgagacatgtgt 792
Db 848 AACGACCTGTCTGGCGTCTCTCTACAGACACAGAGGTGACCTACCTCATGGACATGTC 907
QY 793 caattcgacactgtgttagaacttctgacgctactgaattgtctccattctgtgttctg 852
Db 908 TCCTTCGACACATCTCCACGACGACCGTCGACACCAAGCTGTCCCTCTCTGTGACCTG 967
QY 853 ttcactcagcgaattggatccaatcagactacttgcacagcttgcacagcttgcacagcttgcacag 912
Db 968 TTCCTATGACGAATGGATCACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 1027
QY 913 tacggtgtgtgtaaccattgtgttccagctcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1028 CATGCGCGGTACCGCTCGGCCCGCCACCGAGGCGGTGCGGTACGCTACGCTACGAGCTCATC 1087
QY 973 gctgattgactcactctccagtttcaagacacacacttctactaaccacacttctgacct 1032
Db 1088 GCCGCTCTCAACCTACCTGCTCCAGATGACACACAGCTCCAACACACATTCGACTCT 1147
QY 1033 aaccagctacttcccatggaagctacttgaagctacttgaagctacttctcagcaaacact 1092
Db 1148 AACTCGGCTACCTTTCGCTCACTCTCTCTACGCGGACTTTTCCACAGATAACGGC 1207
QY 1093 atgatatctatttcttctgt 1152
Db 1208 ATCATCTCTATCTCTTGT 1267
QY 1153 tctgttgaactattgaagaactgacggttactctgtcttctgtgtgtgtgtgtgtgtgtgtgt 1212
Db 1268 ACCGTGAGAAATATCCACGACAGATGATTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1327
QY 1213 gctagagcttacgttgaatgacgaatgtcaagctgaaaggaacacattgtgtgtgtgtgt 1272
Db 1328 TCGGCTGT 1387
QY 1273 ttggttaacgacgagattgttccattgcacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1332
Db 1388 TTGGTTAATGATCGCTGT 1447
QY 1333 aagagagacgactcttgaaggtttgtcttctgtctgtgtgtgtgtgtgtgtgtgtgtgtgt 1392
Db 1448 ACCCGAGATAGCTTTGTGAAGGGGTGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1507
QY 1393 tgttctgctta 1403
Db 1508 TGTGTGTGTGT 1518

RESULT 11

AR018076

LOCUS

AR018076 2363 bp DNA

PAT

05-DEC-1998

Sequence 7 from patent US 5780292.

DEFINITION
ACCESSION
AR018076.1 GI:3973679
VERSION
KEYWORDS
SOURCE
ORGANISM

Unknown.
Unclassified.
1 (bases 1 to 2363)
Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
Torckeli,T.K., Cantrell,M., Piddington,C.S., Rambosek,J.A.,
Turunen,M.K. and Fagerstrom,R.B.
Production of phytate degrading enzymes in trichoderma
Patent: US 5780292-A 7 14-JUL-1998;
Location/Qualifiers
1..2363
/organism="unknown"

BASE COUNT 559 a 732 c 510 g 562 t
ORIGIN

Query Match 44.8%; Score 629.4; DB 9; Length 2363;
Best Local Similarity 65.8%; Pred. No. 4.6e-160;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

QY 13 gtcgtgactgcccattgcccacattgctcggttccacactccggtacacgctggtggtcct 72
Db 518 GTGGGACTACTGATCGCTGACAATCTGTGCAGAGTCACTCCGAGCTGGCAGTCCCGCC 577

QY 73 cgtggaattctcaactcttgacactgttacggtgtgtaccagtgtttcccaagaatt 132
Db 578 TCAGAGAAATCAATFCCATCTCCGATACGCTGTCGATCAAGGGTATCAATGCTTCCGAGACT 637

QY 133 tctcaactgtgggtacctactctccatactctctctctctctctctctctctctctctct 192
Db 638 TCGCATCTTTGGGGTCAATACGCGCGGTCTTCTCTCTGGCAAGCAATCGGCCATCTCC 697

QY 193 ccagacgtttccagacgactgtagttagtctctctctctctctctctctctctctctct 252
Db 698 CTTGATGTGCCCGCGGGTTCAGAGTCACTTTCGCTCAGGCTCTCCCGTTCATGGAGCG 757

QY 253 agatacccaactctctcgtcttaaggcttactctgctgttgattgaagctattccaaag 312
Db 758 CGGTATCCGACCGAGTCCAGGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG 817

QY 313 aacgctactgttccaaaggttaagtcagctgttctctctctctctctctctctctctct 372
Db 818 AACGTGACCACTTTGATGGAAATATGCTTCTCCTGAAGACATACAACTACAGCTTGGGT 877

QY 373 gctacacactgactcattcgtgaaacccaaatggttaactctggtattaaagttctac 432
Db 878 GCAGATGACTGACTCCTTCGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937

QY 433 agaagatacaggtgtgggtgctgaaagattgttccattcattagagctctcgtctctgac 492
Db 938 CAGCGATACGAATCGCTCAACAGGAACATCATTTCCGTTTCATCCGATCTCTGGCTCCAGC 997

QY 493 agagtattgctctgctgaaagctcattgaaggtttccaatctgctgaagttgctgac 552
Db 998 CGCGTGTGCGCTCCGGCGAGAAATCATTTGAGGGTTCAGAGCCACCAAGCTGAAGAT 1057

QY 553 ccaggttctcaacacacacacagctctccagttattaaagctgacatcccaagaagatcc 612
Db 1058 CTTGTCGCCAGCGCGGCAATCGTCGCCCAAGATCAGCTGCTGCTATTTCCGAGGCCAGC 1117

QY 613 ggttacacaacacacttgaccacggttaactgactgtcttcgcttcgaacactcgaattagt 672
Db 1118 TCATCCAAACAACACTCGACCCAGGCACTGCTGCTTTTGAACACAGCAAGTTGGCC 1177

QY 673 aacgactgtgaagtaactcactgctgttctgctccagctattagagctagattgaa 732
Db 1178 GATACCGTCGAGCCATTTACCGCCACGCTCGCGCCCTCCATTCGTCAACGCTCTGGAG 1237

QY 733 gctgacttgcaggtgttaacttctgactgacgagacgctgttcttaactgtgacatgtgt 792

Db 1238 AACGACCTGTCTGGCGTGACTCTCAGACACAGAAAGTACCTACTCTATGACATGTGC 1297

QY 793 ccattcgacactgctgcgtagaacttctgacgctactgaaattgtctccattctgtgtttg 852

Db 1298 TCCTTTCGACACCATCTCCACCAGCAGCGTCGACACCAAGCTTCCCGCTTCTGTGACCTG 1357

QY 853 ttoactcacgacgaatggatcccaatcacgactactgtcaaaagcttgggtaagtactacggt 912

Db 1358 TTCAACCATGACGAATGGATCCACTACGACTACCTCCAGTCCCTGAAANAATACATACGCG 1417

QY 913 tacggtctgtagtaccactgggtccagctcaaggtgtgttggtttcgtctcaacgaattgatt 972

Db 1418 CATGGCGAGGTACACCGCTCGCGCCGACCCAGCGCTGGCTACGCTAACGAGCTCATC 1477

QY 973 gctagattgactcaactctccagttccagttccagacacactcttactaaccacactttgactct 1032

Db 1478 GCCGCTCTCAACCCACTCGCCTTCCACGATGACACCAAGCTCCAAACACACCTTGGACTCG 1537

QY 1033 aaccagactacttccactgaaagctactttgtacgctgaactctctcacgacacact 1092

Db 1538 AACCCAGACTACTTCCCGCTCAACTCTCTCTACCGGAGCTTTTCCACGATACAGGC 1597

QY 1093 atgatatactattctctcgtgttggtttgtacaaggttaccagaacattgctactact 1152

Db 1598 ATCATCTCTATCTCTTTGCTTTGGGTCTGTGTACAAACGGCACTAAGCGCTGTCTACACG 1657

QY 1153 tctgtgaactattgaagaactgacggttactctctctctctctctctctctctctctctct 1212

Db 1658 ACCGTGGAGAAATATACCCAGACAGATGAGGTCTCTGCTCTCTGCTGGAGGTTCCGTTGCT 1717

QY 1213 gctagagcttactgtgaaatgatcaatgtcaagctgaaagaaacacattggttagagtt 1272

Db 1718 TCGCGTCTGTAGCTGAGATGATGACAGTGCAGCGCCAGCAGGAGCGCTGCGCTGTC 1777

QY 1273 ttggttaagacagagattgttccattgacaggttctgactgttgaacaggttgggtagatgt 1332

Db 1778 TTGGTTAATGATGCGGTTGTCCCGCTGCATGGGTGTCATTTGATGCTTTGGGAGATGT 1837

QY 1333 aagagagacgactctgttgaaggtttgttctctcgtctagatctggttaactgggtgaa 1392

Db 1838 ACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGTAGATCTGGGGGTGATTTGGCGGAG 1897

QY 1393 tgtttcgctta 1403

Db 1898 TGTTCGTCTTA 1908

RESULT 12

AR051916 2363 bp DNA PAT 29-SEP-1999

LOCUS Sequence 7 from patent US 5830733.

DEFINITION AR051916

ACCESSION AR051916

VERSION AR051916.1 GI:5975280

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2363)

AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
Torckeli,T.K., Cantrell,M., Piddington,C.S., Rambosek,J.A.,
Turunen,M.K., Fagerstrom,R.B. and Houston,C.S.
Nucleic acid molecules encoding phytase and pH2.5 acid phosphatase
Patent: US 5830733-A 7 03-NOV-1998;
Location/Qualifiers
1..2363
/organism="unknown"

BASE COUNT 559 a 732 c 510 g 562 t

ORIGIN

Query Match 44.8%; Score 629.4; DB 9; Length 2363;
Best Local Similarity 65.8%; Pred. No. 4.6e-160;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

QY 13 gtcgtgactgcccattgcccacattgctcggttccacactccggtacacgctggtggtcct 72
Db 518 GTGGGACTACTGATCGCTGACAATCTGTGCAGAGTCACTCCGAGCTGGCAGTCCCGCC 577

QY 73 cgtggaattctcaactcttgacactgttacggtgtgtaccagtgtttcccaagaatt 132
Db 578 TCAGAGAAATCAATFCCATCTCCGATACGCTGTCGATCAAGGGTATCAATGCTTCCGAGACT 637

QY 133 tctcaactgtgggtacctactctccatactctctctctctctctctctctctctctctct 192
Db 638 TCGCATCTTTGGGGTCAATACGCGCGGTCTTCTCTCTGGCAAGCAATCGGCCATCTCC 697

QY 193 ccagacgtttccagacgactgtagttagtctctctctctctctctctctctctctctct 252
Db 698 CTTGATGTGCCCGCGGGTTCAGAGTCACTTTCGCTCAGGCTCTCCCGTTCATGGAGCG 757

QY 253 agatacccaactctctcgtcttaaggcttactctgctgttgattgaagctattccaaag 312
Db 758 CGGTATCCGACCGAGTCCAGGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG 817

QY 313 aacgctactgttccaaaggttaagtcagctgttctctctctctctctctctctctctct 372
Db 818 AACGTGACCACTTTGATGGAAATATGCTTCTCCTGAAGACATACAACTACAGCTTGGGT 877

QY 373 gctacacactgactcattcgtgaaacccaaatggttaactctggtattaaagttctac 432
Db 878 GCAGATGACTGACTCCTTCGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937

QY 433 agaagatacaggtgtgggtgctgaaagattgttccattcattagagctctcgtctctgac 492
Db 938 CAGCGATACGAATCGCTCAACAGGAACATCATTTCCGTTTCATCCGATCTCTGGCTCCAGC 997

QY 493 agagtattgctctgctgaaagctcattgaaggtttccaatctgctgaagttgctgac 552
Db 998 CGCGTGTGCGCTCCGGCGAGAAATCATTTGAGGGTTCAGAGCCACCAAGCTGAAGAT 1057

QY 553 ccaggttctcaacacacacacagctctccagttattaaagctgacatcccaagaagatcc 612
Db 1058 CTTGTCGCCAGCGCGGCAATCGTCGCCCAAGATCAGCTGCTGCTATTTCCGAGGCCAGC 1117

QY 613 ggttacacaacacacttgaccacggttaactgactgtcttcgcttcgaacactcgaattagt 672
Db 1118 TCATCCAAACAACACTCGACCCAGGCACTGCTGCTTTTGAACACAGCAAGTTGGCC 1177

QY 673 aacgactgtgaagtaactcactgctgttctgctccagctattagagctagattgaa 732
Db 1178 GATACCGTCGAGCCATTTACCGCCACGCTCGCGCCCTCCATTCGTCAACGCTCTGGAG 1237

QY 733 gctgacttgcaggtgttaacttctgactgacgagacgctgttcttaactgtgacatgtgt 792

Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

```
QY 13 gctgctgactgtccatccacccctgttgcgttccacatccggtaccgcttggtgct 72
Db 518 GTGGGACTACTGATCGCTGACAACTGTGCGAGATCACTCCGAGCTGGCAGTCCCGCC 577
QY 73 cgtggaatttcactcttctgtgacactgttgacggtgtgtaccgaatttccagaatt 132
Db 578 TCGAGAAATCAATCCACTTGGGATACGCTGATCAAGGGTATCAATGCTTCTCGAGACT 637
QY 133 tctcaattgtgggtaccatctccatctcttcttggcagacgaattctgtattct 192
Db 638 TCGCATCTTTGGGGTCAATACCGCCGCTTCTCTCTGCGAAAGAAATCGGCGATCTCC 697
QY 193 ccagactgtccagacgactgtagagttacttctgttccaaagtgttctagacaggtgt 252
Db 698 CCGTGTGTCGCGCGGTTGCAGAGTCACTTTCCTCAGGCTCTCTCCCGTCATGGAGCG 757
QY 253 agataccaaactcttctgtgttaaggttactctgttctgttgaagctattccaaag 312
Db 758 CGGTATCCGACCGAGTCCAAAGGCAAGAAATACTCCGCTCTCATTCAGGAGATCCAGCAG 817
QY 313 aacgtactgtttcaaggttaagtaagctgttcttcttgaagacttacaactttgggt 372
Db 818 AACGTGACCACTTTGATGGAAATATGCTTCTCTGGAAGACATACAACTACAGCTTGGGT 877
QY 373 gctgacgactgtactccatctcgtgtaaaacaaatgttaactctgttattgaattctac 432
Db 878 GCAGATGACCTGACTCCTTCGAGAGCAGGAGTACTGACCTCCGCGATCAAGTTCTAC 937
QY 433 agaagatacgaagcttgggtgagaagattgttccattcattagagcttctgttctgac 492
Db 938 CAGCGATACGAATCGCTACAAAGGAACATCATTCCTGTCATCGATCTCTGCTCCAGC 997
QY 493 agagtattctctgtctgaaagtctgaaagtctgaaagtcttccactctgtaagttgctgac 552
Db 998 CGGTGTGTCGCTCCGCGGAGAAATTCATGTAGGGCTTCCAGAGCACCAGGCTGAAGGAT 1057
QY 553 ccaggttctcaacacacacacagcttccaggttattacagctgatactccagaagatcc 612
Db 1058 CTTGTCGCCAGCGCGGCAATCGTCGCCCAAGATCGACGTGTCATTTCCGAGCCAGC 1117
QY 613 ggttaacaaacacttgcacacgcttactgtactgtcttgcagactctgtaattaggt 672
Db 1118 TCATCAACAAACACTCTCGACCCAGGACCTGCACTGTCTTTGAAGACAGCGAATTGGCC 1177
QY 673 gacgacttgaagctaaactctactgcttctgttctccagctattagagctagattgaa 732
Db 1178 GATACGCTGGAAGCCAAATTCACCGCCACGTTCCGCCCTCCATTCGTCAACGCTGGAG 1237
QY 733 gctgacttgcaggtgttacttgcagacgagacgtgttacttgcagacatgtgt 792
Db 1238 AACGACTGTCTGGCGTGACTCTCAGACACAGAAAGTACCTGATGATGACATGTGC 1297
QY 793 ccattgcagactgtgctagaactctgacgctactgactgacttccattctgtctgttg 852
Db 1298 TCCTTCGACACCATCTCCACAGACCGCTGCGACACCAAGCTGTCCCTCTCTGTGACCTG 1357
QY 853 ttcactcagcagaatgatccaatcacactacttgcagaaagtgttgtaagtaactacgt 912
Db 1358 TTCACCATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1417
QY 913 tacgtgtgttaacccattgggtccagctcaagctcaaggtgtgttgcgaagaattgatt 972
Db 1418 CATGGCGAGGTAAACCGCTCGGCCCGGACCCAGGCGCTGCGCTAGGCTAAGGAGTATC 1477
QY 973 gtagattgactcactctccagttcaagaccacacttctactaacacactttgactct 1032
Db 1478 GCGCGCTCTCACCACTCGCCTGTCCACGATGACACCACTCTCAACCACTCTGACTTCG 1537
QY 1033 aaccagactacttctccattgaagacttctgtacgtgacttctctcagcaaacact 1092
Db 1538 AACCCAGTACTCTCCCGCTCAACTCTACTCTACGGGGACTTTTCCCAAGATTAACGGC 1597
```

```
QY 1093 atgatctatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1152
Db 1598 ATCATCTCTATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1657
QY 1153 tctgtgaattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1212
Db 1658 ACCGTGGAGAAATATACACCCAGACAGATGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1717
QY 1213 gctagagcttgcgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1272
Db 1718 TCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1777
QY 1273 ttggttaacagacagagttgttccattgcacggtgtgctgttgcagaaagtgtggttagatgt 1332
Db 1778 TTGGTTAATGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1837
QY 1333 aagagagacgacttgcgttgaaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1392
Db 1838 ACCCGGATGATGCTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGCGGAG 1897
QY 1393 tgttctcgttta 1403
Db 1898 TGTCTCTGCTTA 1908

RESULT 13
AR053934
LOCUS AR053934 2379 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5834286.
ACCESSION AR053934
VERSION AR053934.1 GI:5978796
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2379)
AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Fagerstrom,R.B.,
Miettinen-Oinonen,A.S.K., Turunen,M.K., Rambosk,J.A.,
Piddington,C.S., Houston,C.S. and Cantrell,M.A.
Recombinant cells that express phytate degrading enzymes in desired
ratios
TITLE JOURNAL
FEATURES Patent: US 5834286-A 1 10-NOV-1998;
Location/Qualifiers
1..2379
BASE COUNT 561 a 735 c 518 g 565 t
ORIGIN

Query Match 44.88; Score 629.4; DB 9; Length 2379;
Best Local Similarity 65.88; Pred. No. 4.6e-160;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

QY 13 gctgctgactgtccatccacccctgttgcgttccacatccggtaccgcttggtgct 72
Db 534 GTGGGACTACTGATCGCTGACAACTGTGCGAGATCACTCCGAGCTGGCAGTCCCGCC 593
QY 73 cgtggaatttcactcttctgtgacactgttgacggtgtgtaccgaatttccagaatt 132
Db 594 TCGAGAAATCAATCCACTTGGGATACGCTGATCAAGGGTATCAATGCTTCTCGAGACT 653
QY 133 tctcaattgtgggtaccatctccatctcttcttggcagacgaattctgtattct 192
Db 654 TCGCATCTTTGGGGTCAATACCGCCGCTTCTCTCTGCGAAAGAAATCGGCGATCTCC 713
QY 193 ccagactgtccagacgactgtagagttacttctgttccaaagtgttctagacaggtgt 252
Db 714 CCGTGTGTCGCGCGGTTGCAGAGTCACTTTCCTCAGGCTCTCTCTCTCTCTCTCTCTCTCT 773
QY 253 agataccaaactcttctgtgttaaggttactctgttctgttgaagctattccaaag 312
Db 774 CGGTATCCGACCGAGTCCAAAGGCAAGAAATACTCCGCTCTCTCTCTCTCTCTCTCTCTCT 833
```

QY 313 aacgctactgcttcaaggtaagtaagctgttcttgaagaacttaacactacactttgggt 372
 DB 834 AACCTGACCACTTTGATGTAATAATGCTTCTCTGAAGACATACAACTACAGCTTGGGT 893
 QY 373 gctgaacactgactccattcggtgaaacaaactaactgtaactctggtattaaagtctac 432
 DB 894 GCATGACCTGACTCCCTTCGGAGACGACGAGCTAGTCACTCCGGCATCAAGTTCTTAC 953
 QY 433 agaagatacaaggcttggcttagaagaattgttccattcattagattcattggttctgac 492
 DB 954 CAGCATACGAATCGCTACAAGGAACATCATTCCTGTTTCACTCCGATCTCTGGCTCCAGC 1013
 QY 493 agagtattgtctgtcgtgaaagactcattgaaggtttccaaactcgttaagtggctgac 552
 DB 1014 CGCTGTGCTCGCTCCGGCGAATAATTCATTGAGGGCTTCCAGAGACCAAGCTCAAGGAT 1073
 QY 553 caaggttctcaaccac 612
 DB 1074 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
 QY 613 ggttacaac 672
 DB 1134 TCATCCAAACACTCTCGACCAAGGACCTGCACTGTCTTTGAAGACACGCAATTTGCC 1193
 QY 673 gacgactgtaagctaaactcactccttctgctcactcagctattagcttagtgaa 732
 DB 1194 GATACCTGCAAGGCAATTTCCACCGCACAGTTCGCCCTCCATTCGTCACGCTCGAG 1253
 QY 733 gctgacttgcacaggttgaacttgaactgacgaagacgttgttacttgatggacatggt 792
 DB 1254 AACGACCTGTCTGGGTGACTCTCACAGACACAGAACTGACCTACCTCATGGACATGTC 1313
 QY 793 ccattcgacactgctgtagaactctgacactgactgaattgttccactctggtttg 852
 DB 1314 TCCTTCGACACCATCTCCACCAAGCCTCGACACCAAGCTGTCCCGCTTCTGTGACCTG 1373
 QY 853 ttcaactcagcgaatgatccaatcagactacttgaaagcttgggttaagtactacgtt 912
 DB 1374 TTCACCATGACGAATGATCCACTACGACACTACCTCCAGTCCCTGAAATACTACGGC 1433
 QY 913 tacggtgtgtgtaacccattgggttcagctcaagtggttggtttgcgttaacgaattgatt 972
 DB 1434 CATGGCGCAGGTAAACCGCTCGCCCGCAGCCAGGCGCTGCGCTAGCTTAACGAGCTATC 1493
 QY 973 gctagattgactcactctccagttcaagacacacacttctactaaccacactttgactct 1032
 DB 1494 GCGCGTCTCACCACTCGCCTGTCCACGATGACACCACTCCAGTCCCTGAAATACTACGGC 1553
 QY 1033 aaccagactacttccattgaacgtacttctgtacgctgacttctctcagcaaacact 1092
 DB 1554 AACCCAGCTACCTTCGCGTCACTCTACTCTCTACGCGGACTTTTCCACGATAACGGC 1613
 QY 1093 atgatatctattcttctgcttgggttgaacacggttcaaacgacttcttactact 1152
 DB 1614 ATCATCTCTATCTCTTCTTGGTCTGTACACGCGACTTAACCGCTGTGCTTACCACG 1673
 QY 1153 tctgttgaactctatgaagaactcaggttactctgcttcttggactgttccattcgt 1212
 DB 1674 ACCGTGGGAATATACCCAGACAGATGGGTCTCTGCTGTGGACGTTCCGTTTGTCT 1733
 QY 1213 gctagagcttacttgaatgatgcaatgtcaagctgaaaggaacacttggtagagatt 1272
 DB 1734 TCGCGTCTGTACGTCGAGATGATGATGCCAGGCGGACGAGGAGCGGTGCTGCTGTC 1793
 QY 1273 ttggttaacagacagagatttccattgacggttctgctgttgaagaagtgtgggttagatt 1332
 DB 1794 TTGGTTAATGATGCGGCTGTCCCGCTGCTGATGGGTGTCGAATGATGCTTTGGGAGATGT 1853
 QY 1333 aagagagacgacttctgtaaggttcttctcgttagatctggttgtaactgggtgaa 1392
 DB 1854 ACCGGGATAGCTTTGTGAGGGGGTTGACCTTTGCTAGATCTTGGGGTGTATTTGGCGGAG 1913

QY 1393 tggctcgctta 1403
 DB 1914 TGTCTGCTTA 1924
 RESULT 14
 ASNPHTAS 2379 bp DNA PLN 23-MAR-1994
 LOCUS Aspergillus niger var awamori phytase gene, complete cds.
 DEFINITION L02421
 ACCESSION L02421
 VERSION L02421.1 GI:166518
 KEYWORDS phytase.
 SOURCE Aspergillus niger (strain ALK0243, sub-species awamori) DNA.
 ORGANISM Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 1 (sites)
 Piddington, C.S., Houston, C.S., Paloheimo, M., Cantrell, M.,
 Miettinen-Oinonen, A., Nevalainen, H., and Rambosek, J.
 The cloning and sequencing of the genes encoding phytase (phy) and
 pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
 awamori
 Gene 133 (1), 55-62 (1993)
 94040796
 2 (bases 1 to 2379)
 Carter, J.R., Franden, M.A., Aebersold, R.H. and McHenry, C.S.
 Molecular cloning, sequencing and overexpression of the gene
 encoding the psi subunit of E. coli DNA polymerase III holoenzyme
 Unpublished (1992)
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1..2379
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 420..463
 /citation=[2]
 /evidence=experimental
 join(420..463,566..1925)
 /EC_number="3.1.3.8"
 /citation=[2]
 /codon_start=1
 /function="conversion of phytate to inositol and inorganic
 phosphorus"
 /evidence=experimental
 /product="phytase"
 /protein_id="AA16898.1"
 /db_xref="GI:166519"
 /translation="MGVSAVLPLLYLLAGVTSGSLVAPSRNOSTCDTVDQYQCFSET
 SHLQYAPFFSLANESATSPDVPAGCRVTFQAQLSRHGARYPTESKGYKYSALIEI
 QONVTFDQKYAFATYNSLGADDLTPFGEOLVNSGIFKVFYKESLRNIIPIRS
 SSSRVIASGEKFIQFQSTKLKDPRAOPGSSPKIDVVISSSSNNTLDPTCTVF
 EDESLADTVEANFTAFAPSRQRLENLDSVTLTDEVTYVMDMCSFTDITSTPVD
 KISPFCDLETHDEWIHYDLQSKKYYHGAGNPLGPTGGVGMAYANELLARLTHSHVD
 DTSNNHTLDSNPATPLNSTLYADFSDNGIISILFALGLYNGTAPLSTIVENTIQ
 DGFSSAWTVFFASRLYVENMQCAEQEPLRVLRVNDRVVPLHGPCIDALGRCTDSFV
 RGLSFARSGGDWAECSA"
 566..1925
 /note="putative"
 BASE COUNT 561 a 735 c 518 g 565 t
 ORIGIN
 Query Match 44.8%; Score 629.4; DB 13; Length 2379;
 Best Local Similarity 65.8%; Pred. No. 4.6e-160;
 Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
 QY 13 gtcgtgctactgtccattgcccactgttgcgttccacatcccggtaccgcttgggtcct 72
 DB 534 GTGGGACTACTGATCGTACAACTCTGTGCAGAGTCACTCCGGAAGTGGCAGTCCCGCC 593
 QY 73 cgtggttaattctactcttctgacactcttgacggtgtgttaccacatgttcccgagaaatt 132
 DB 594 TCGAGAAATCAATCCACTTGGGATACGATACGATCAAGGGTATCAATCTCTCCGAGACT 653

QY 133 tctcactgttggtggtacactactcctccatactctctcttcttggcagacgaatctgctattctt 192
 Db 654 TCGCATCTTTGGGGTCAATACGCGCGCTCTCTCTCTGGCAACGAATCGCCATCTCC 713
 QY 193 ccagacttcagacagactgtagacttacttctgctcaagtcttctgtagacacaggtgct 252
 Db 714 CCTGATGTGCGCGCGGTGCGAGAGTCACTTTCCTCAGGTCCTCTCCGTCATGGAGCG 773
 QY 253 agatacccaacttctctgctcctaaggtctactctctctctgattgaagctattcaaaaag 312
 Db 774 CGGTATCCGACCGAGTCCCAAGGGAAGAAATACTCCGCTCTCAATGAGGAGATCCACGAG 833
 QY 313 aacgtactgtttccaaaggttaagctcttcttctgattgaagctattcaaaaag 372
 Db 834 AAGGTGACCACTTTGATGGAANAATGCTTCTCTGAAGACATACAACTACAGCTGGGT 893
 QY 373 gctgacgactgactcctcattcgtgaaacacaaatggttaactctggtatttaagttctac 432
 Db 894 GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGTACTCAACTCCGCGCATCAAGTTCTAC 953
 QY 433 agaagatacaaggtcttggctagaagaattgttccattcattagagcttctggttctgac 492
 Db 954 CAGCGATACGATCGCTCAAGAGGAACATCATTCGCTTCATCGATCCTCTGGCTCCAGC 1013
 QY 493 agagtattgtctgtgaaagtctcattgaaggtttccattcctcctgaagttggtgac 552
 Db 1014 CGCGTGATCGCTCCGCGGAGAAATTCATTGAGGGCTTCAGAGCACCAGCTGAAGGAT 1073
 QY 553 ccaggttctcaac 612
 Db 1074 CCGTGGCCGACCGCGGCCACTCGTCCGCAAGATCGACGTGCTCATTTCCGAGGCCAGC 1133
 QY 613 ggttac 672
 Db 1134 TCATCCAAACACTCTCGACCGACGACCTGCACCTGCTTTGAAGACAGCGAATGGGC 1193
 QY 673 gacgacttgaagactaactcactgttctgctcctcactgcttgaagactgaattaggt 732
 Db 1194 GATACCGTGAAGCAATTTTACCGCGACGCTGCGCCCTCCATTCGTCACGCTCGGAG 1253
 QY 733 gctgacttgcaggtgttacttgcagacagacagctgttactgattgacatgtgt 792
 Db 1254 AACGACCTGTCTGGCTGACTCTCAGACACAGAGAAGTGAACCTACCTCATGACATGTGC 1313
 QY 793 ccattgacactgtcgtagacacttctgacgtactgactgaattgtctcattctgtgcttg 852
 Db 1314 TCCTTGACACCATCTCCACGACCGCTGACACACCAAGCTGTCCCTCTCTGTGACCTG 1373
 QY 853 ttcaactcagcagaaatggatccaatacagactacttgcagagcttgcagagcttgcagct 912
 Db 1374 TTCACCATGACGAATGGATTCACACTACGACTACCTCCAGTCCCTGAAAAAATACTACGGC 1433
 QY 913 tacggtgtggttaacccattggttccagctcagctcaggtgttgcctgaagcgaattgatt 972
 Db 1434 CATGGCGCAGGTAACCGCTCGCGCCGACCCAGGCGGTGGGTACGCTAACGAGCTCATC 1493
 QY 973 gctagattgactcactctccagttcaagacacacacttactaaccacacacttggactct 1032
 Db 1494 GCGCGTCTCACCCACTCGGCTCTCCACGATGACACACAGCTCCCAACCAACCTTGACCTG 1553
 QY 1033 aacccagctacttcccaactgacgactcttctgacgtgacttctctcagacacacact 1092
 Db 1554 AACCCAGTACCTTCCGCTCAACTCTACTCTACTACGCGGACTTTTCCACGATAAACGGC 1613
 QY 1093 atgatattcttcttctgcttgggttgggttgcacacggtacacacacacacacacacacac 1152
 Db 1614 ATCATCTCTACTCTTGGCTTGTACACGCGCATTAAGCCCTGTCTACACG 1673
 QY 1153 tctgttgaactctattgaagaactgaaggttactctgcttcttctgactgttccattgct 1212
 Db 1674 ACCGTGGAGATATACCCACGACAGATGGGTCTCTGCTGCTTGGACGGTTCCGCTTGTCT 1733

QY 1213 gctagagcttacgttgaatgatgcaatgtcaagctgaaagagaaacacattggttagagtt 1272
 Db 1734 TCCGCTGTGACTGAGATGATGCAAGTCCAGGCGGACGAGCGCTGCTGCTCGCTGC 1793
 QY 1273 ttggttaacagacagagttgttccattgacagctgtgtgctgttgacaaagtgtggttagatgt 1332
 Db 1794 TTGTTTAATCATCGCTTGTCCCGCTGCATGGGTGTCCTCAATTGATGCTTTGGGAGATGT 1853
 QY 1333 aagagagacgactcttgaaggttcttctgctgactgactgagttggtgaactgggtgaa 1392
 Db 1854 ACCCGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTTGGGGTGATTTGGCGGAG 1913
 QY 1393 tgttctgctta 1403
 Db 1914 TGTCTGCTTA 1924

RESULT 15
 AB022700 1515 bp DNA PLN 26-JAN-1999
 DEFINITION Aspergillus niger gene for phytase, complete cds.
 ACCESSION AB022700
 VERSION AB022700.1 GI:4185609
 KEYWORDS phytase.
 SOURCE Aspergillus niger (strain:SK-57) DNA.
 ORGANISM Aspergillus niger
 Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 1 (bases 1 to 1515)
 Nagashima,T., Kondo,H., Anazawa,H. and Terasaki,Y.
 Phytase having high-affinity for phytic acid
 Published Only in Database (1999) In press
 2 (bases 1 to 1515)
 Nagashima,T., Kondo,H., Anazawa,H. and Terasaki,Y.
 Direct Submission
 Submitted (21-JAN-1999) to the DDBJ/EMBL/GenBank databases. Tadashi
 Nagashima, Shin Nihon Chemical Co.,Ltd, Research and Development
 Div.; Showa-cho 19-10, Anjo, Aichi 446-0063, Japan
 (E-mail:nagashima@dl.dion.ne.jp, Tel:81-566-76-5171(ex.264),
 Fax:81-566-75-0010)

FEATURES
 Location/Qualifiers
 1..1515
 /organism="Aspergillus niger"
 /strain="SK-57"
 /db_xref="taxon:5061"
 join(1..44,156..1515)
 /gene="phyA"
 join(1..44,156..1515)
 /gene="phyA"
 /codon_start=1
 /product="phytase"
 /protein_id="BAA74433.1"
 /db_xref="GI:4185610"
 /translation="MGVSALLPLVLLSGVTSLAVPASRNSTCDTVDOGYQCRSET
 SHLGOVAPFSLANKSAISPDVPACGHVTFQVLSRHGARIPTDSKKAKYALIEEI
 QONATTEKYALFKTYNSLQADLTPEGLVNSGVKVFORYESTLRNTVPTIRS
 SGRSRVIAQGNKFIQEGTQKLKDPQSQSPKIDVVISEASTNNTLDPTCTVF
 EDELADDEANFTATTFVSIQRLENDLSGSLDTEVYILMDSCFTIISTVDT
 KLSPPDLFTHEWINDYLIQSLNKYHGAGNPLCPTOGVYANELIARLTHSPVHD
 TYSNHTLDSNPATFPLNSTLYADFSHDNGIISILFALGLYNGTKPLSSTTAENITQT
 DGFSSAWTVPFASRMVYEMMOCQSEBPFLRVLVNDRVPLVHGGCPDVALGRCTROSFV
 KGLSFARSGDNGECFA"

BASE COUNT 317 a 478 c 365 g 355 t
 ORIGIN

Query Match 43.6%; Score 612.4; DB 12; Length 1515;
 Best Local Similarity 65.5%; Pred. No. 1.8e-155;
 Matches 895; Conservative 0; Mismatches 471; Indels 0; Gaps 0;

QY 38 tgtctggttccacatccggtacgccttgggtctctggttgaattctcactcttggaca 97
 Db 149 TGCACAGAGTCACTCCGGACTGGCAGTCCCGCCCTCGAGAAATCAATCCACTTGGGATA 208

QY	1178	acgggttaactctgtctcttctggaactgttccattcgtcgtctagagcttaactgtgaaatgatgc	1237
Db	1289	ATGGGTTCTCATCTGCTGGAGGGTTCCTTTCCGCGTCGCGCATGTACGTCGAGATGATGC	1348
QY	1238	aatgtcáagctgaaagaaacccattgggttagaggttttgggttaacgacagaggttggccat	1297
Db	1349	AATGCCAGTCCGAGCAGGAGCCCTTTGGTCCGTCTTGGTTAATGATCGTGTGTTCGCGC	1408
QY	1298	tgcaaggttggctgtgttgacaagttgggttagatgtaagagagacgacttcgttgaaggtt	1357
Db	1409	TGCATGGCTGTCCGGTTGATGCTTTGGGAAGATGTACGCGGGATAGCTTCGTGGAAGGGT	1468
QY	1358	tgcttctcgtatgactcgtgtgaactgggtgaatgttcttgcctta	1403
Db	1469	TGAGCTTTGCCAGATCTGCGGTGATTGGGGGAGTGTTCGCTTA	1514

Search completed: October 26, 2001, 17:47:03
Job time: 8779 sec

QY	98	ctgttgacggttggttaacaaatgtttccagaaatttctcaactgtgtgggtacactactctc	157
Db	209	CGGTGATCAGGGGTATCAATGCTTCGAGAGACTTCGGCATCTTTGGGGCCAATACGCGC	268
QY	158	catactctcttggcgagacaaactcgtctatttctcagacggttccagacgactgttagag	217
Db	269	CGTCTTTCTCTGCAACAACAAATCGGCCATCTCCCTGTATGTTCTTGCCTGGATGCCATG	328
QY	218	ttaactttggttcaagtgttctagacacggtgctagatacccaactcttctcgtcgtcta	277
Db	329	TCACCTTTTCGCCAGGTTCTTCCCGCCATGAGACGAGCATCCACGACTCCCAAGGSCA	388
QY	278	agcgttactctgtgttgaagtattcaaaagacgctactcgttctcaagggtaagt	337
Db	389	AGAAATACTCCGCTCTCATCGAGGAGATCCAGCAGAACCGCACAACCTTTCGAGGGGAAT	448
QY	338	acgctttttgaagacttaacactacactttgggtgctgacgacttgactccattcgtg	397
Db	449	ATGCCCTTCTGAAGACATACAACTACAGCCTGGGGCGGGATGATCTGACTCCCTTCGGAG	508
QY	398	aaacacaaatgttaactctgttattaaagtctctacagaagatacaaggcttttggctagaa	457
Db	509	ACGAGAGCTGTCAACTCCGGCGTCAAGTCTACACGCGATACGAATCGCTCACAGAA	568
QY	458	agattgttccattcattagagcttctgtgtcttgacagaggttattgctctcgtcgtaaaaagt	517
Db	569	ACATTGTCCGTTTCATCCGATCTCAGGCTCCAGCGCGGTGATTGCCCTTGCCAAATAAAT	628
QY	518	tcattgaaggtttccaactcgttaagtgtgctgacccaggttctcaaccacacaaagctt	577
Db	629	TCATCAGAGGGCTTCCAGAGCACTAAGCTGAAGGATTCCTTCGTGCCCGCCGCGCCAAATCGT	688
QY	578	ctccagttataacgtgatactccagaggaatccggtttacaaacacacttttggaccacy	637
Db	689	CGCCCAAGATCGAGTGGTTCATTTAGAGGCCAGCACATCCAAACACACTCTCGATCCGG	748
QY	638	gtacttgtactcgttctcgaagactcgtgaattaggtgaagacgttgaagacttaactcactg	697
Db	749	GCACCTGCACCGTTTTCGAAGATAGCGAATTTGGCGGATGACATCGAAGSCAAATTCACCG	808
QY	698	cttgtgtcgtccagctattagacttagattggaagctgacttcgaggttacttttga	757
Db	809	CCACGTTCCGTCCTCCATTCGTCAACGCTGGAGAACGACTTGTCTGGCGTGTCTCTCA	868
QY	758	ctgacgaagacggttgttacttgatggacatgtgtccattcgcacactcgtcgttagaactt	817
Db	869	CGGACACAGAAGTGACCTACCTCATGGACATGTGCTCTCGACACCACTCTCCACCAGCA	928
QY	818	ctgacgctactgaattgtctccattctgtgttctgttctcactcagcagatggtccaat	877
Db	929	CCGTCGACACCAAGCTGCCCTTCTGTGACCTGTTCACCCCATGAAGAATGGATCAACT	988
QY	878	acactacttgcgaagcttgggttaagtactacggttacggtgctggttaacccattgggtc	937
Db	989	ACGACTACCTCCAGTCCCTGAAACAATACTACGCCCATGGCGAGCTAACCGCTCGGCC	1048
QY	938	cágtcaaggtgttgggttcgtcaagaaattgattgctagattgactcactctccagtcc	997
Db	1049	CGACCCAGGGCGTCGGCTACGCTAACGAGCTCATCGCCGCTCTCACCCACTCGCCTGTCC	1108
QY	998	aagaccacacttctactaacacacitttgactcactaacccagctactttccattgaagc	1057
Db	1109	ACGATGACACAGCTCCAAACACACATTTGACTCCCAACCCGGCTACTTCCCGCTCACT	1168
QY	1058	ctacttgtacgctgacttctctcagcaacacactatgatatttttcttcgctttgg	1117
Db	1169	CCACTCTCTATGGGACTTTTGGCATGATAACGGCATCATCTCTATCTCTTTTGTCTTGG	1228
QY	1118	gtttgtacaaoggttaccacccattgtctactacttctgttgaatctatttgaagaactg	1177
Db	1229	GTCTGTACACAGCGGCCAAGCCGCTGTCTTCCACGACCGGGAGATATCACCCAGACCG	1288

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1404	100.0	1404	20	AA931520	Consensus phytase-	
2	1404	100.0	1404	21	AA973233	Consensus phytase	
3	1404	100.0	1404	21	AA957915	DNA encoding a mut	
4	1388	98.9	1404	21	AA973291	Consensus phytase	
5	1384.8	98.6	1404	21	AA973290	Consensus phytase	
6	1360.8	96.9	1426	20	AA927423	Ascomycete consens	
7	1360.8	96.9	1426	20	AA923022	Fungal phytase gen	
8	1360.8	96.9	1426	21	AA973231	Consensus phytase	
9	1360.8	96.9	1426	21	AA959637	Consensus phytase	
10	1327.2	94.5	1404	20	AA931521	DNA encoding phyt	
11	1327.2	94.5	1404	21	AA973234	Consensus phytase	

- 1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

SUMMARIES

XX
P-PSDB; AAY4

DNA encoding a mut
Consensus phytase
Consensus phytase
Consensus phytase
DNA encoding phyt
Consensus phytase
Consensus phytase
DNA encoding phyt
Aspergillus ficum
Chromosomal phyt
A. ficum phytase
Phytase gene. Asp
Sequence, from ove
Aspergillus ficum
A. fumigatus phyt
Aspergillus fumiga
DNA encoding a mut
Phytase gene. Asp
A. niger phytase g
Aspergillus niger
Aspergillus fumiga
A. fumigatus phyt
Aspergillus nidula
A. nidulans phytic
Phytase chimeric g
Phytase thermo
Talaromyces ph
Aspergillus terre
Phytase gene. Asp
A. terreus phytase
P. hordei phytase
P. hordei phytase
Part of P. hordei
P. hordei phytase
P. hordei phytase

PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 9; 7lpp; English.

XX This sequence encodes the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.
 XX Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;

Query Match 100.0%; Score 1404; DB 20; Length 1404;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcggttctgctgactgctccattgccaacttgggttccacatccgggtacc 60
 Db 1 atggcggttctgctgactgctccattgccaacttgggttccacatccgggtacc 60

QY 61 gcoctgggtctgctggttaattctcaacttctgtgacactgttgacagtggttaccatgt 120
 Db 61 gcoctgggtctgctggttaattctcaacttctgtgacactgttgacagtggttaccatgt 120

QY 121 ttccagaattctcaacttgggtgactactctccatactctcttggcagacgaa 180
 Db 121 ttccagaattctcaacttgggtgactactctccatactctcttggcagacgaa 180

QY 181 tctgctatttctccagacgttccagacgactgtagattacttctgtaagtttggct 240
 Db 181 tctgctatttctccagacgttccagacgactgtagattacttctgtaagtttggct 240

QY 241 agacacggtgctagataccacacttcttggctgctaaagcttactctgcttgattgaa 300
 Db 241 agacacggtgctagataccacacttcttggctgctaaagcttactctgcttgattgaa 300

QY 301 gctattcaaaagacgctactgcttccaaaggttaagtaagcttcttgaagacttacac 360
 Db 301 gctattcaaaagacgctactgcttccaaaggttaagtaagcttcttgaagacttacac 360

QY 361 tacacttgggtgctgacgactgactccattgctgaaacacaaatggttaactctgt 420
 Db 361 tacacttgggtgctgacgactgactccattgctgaaacacaaatggttaactctgt 420

QY 421 attaaagttcacagaagatcacaggcttggctagaaagattgtccattcattagagct 480
 Db 421 attaaagttcacagaagatcacaggcttggctagaaagattgtccattcattagagct 480

QY 481 tctggttctgacagattattgcttctgctgaaagttcattgaaagtttccaatctgt 540
 Db 481 tctggttctgacagattattgcttctgctgaaagttcattgaaagtttccaatctgt 540

QY 541 aagtgtgctgaccaggttctccacacacacacagcttccagtttattacgtgacatt 600
 Db 541 aagtgtgctgaccaggttctccacacacacacagcttccagtttattacgtgacatt 600

QY 601 ccagaagatccggttacacaacacttggaccaggttactgtactgtcttcgaagac 660
 Db 601 ccagaagatccggttacacaacacttggaccaggttactgtactgtcttcgaagac 660

QY 661 tctgaattagtgacagcttgaagcttaacttcaactgttcttgcgtccagctattaga 720
 Db 661 tctgaattagtgacagcttgaagcttaacttcaactgttcttgcgtccagctattaga 720

QY 721 gctagattgaagctgacttgcaggttacttgaactaacacagacgttgttacttg 780
 Db 721 gctagattgaagctgacttgcaggttacttgaactaacacagacgttgttacttg 780

QY 781 atggacatgtgtcattcagacactgctcgtagaaacttctgacgctacigaaattgtctcca 840
 Db 781 atggacatgtgtcattcagacactgctcgtagaaacttctgacgctacigaaattgtctcca 840

QY 841 tctgtgcttgggttccactcacgacgaatggtatccaatcagactacttgcgaagcttgggt 900
 Db 841 tctgtgcttgggttccactcacgacgaatggtatccaatcagactacttgcgaagcttgggt 900

QY 901 aagtaactacggttacggtgctggttaacccattgggtccagctcaaggttgggttctcgt 960
 Db 901 aagtaactacggttacggtgctggttaacccattgggtccagctcaaggttgggttctcgt 960

QY 961 aacgaattgattgctagattgactcactctccagttccaaagacacacttctactaacac 1020
 Db 961 aacgaattgattgctagattgactcactctccagttccaaagacacacttctactaacac 1020

QY 1021 acttggacttaacccagctacttccattgaaacgctacttctgacgctgacttctct 1080
 Db 1021 acttggacttaacccagctacttccattgaaacgctacttctgacgctgacttctct 1080

QY 1081 cagacacacactatgatatctatttctgcttgggtttgtacaacggtaccacagcca 1140
 Db 1081 cagacacacactatgatatctatttctgcttgggtttgtacaacggtaccacagcca 1140

QY 1141 ttgtactacttctgttgactctatttgaagaactgacggttactctgtcttctggact 1200
 Db 1141 ttgtactacttctgttgactctatttgaagaactgacggttactctgtcttctggact 1200

QY 1201 gttccattcgtcgtacagcttactggttgaatgatgcaatgtcaagctgaaaggaacca 1260
 Db 1201 gttccattcgtcgtacagcttactggttgaatgatgcaatgtcaagctgaaaggaacca 1260

QY 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttggctgttgacaag 1320
 Db 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttggctgttgacaag 1320

QY 1321 ttggttagagttgaagagacgacttctggtgaaggttcttcttcgtaagcttgggt 1380
 Db 1321 ttggttagagttgaagagacgacttctggtgaaggttcttcttcgtaagcttgggt 1380

QY 1381 aactgggtgactgttctcgttaa 1404
 Db 1381 aactgggtgactgttctcgttaa 1404

RESULT 2

AAA73233

ID AAA73233 standard; DNA; 1404 BP.

AC AAA73233;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 1 thermo 8 q50t, k91a polynucleotide SEQ ID NO:28.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 XX temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure; ds.

OS Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK000025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX

PI Lehmann M;

XX WPI: 2000-491161/43.

DR P-PSDB; AAB20526.

XX

Novel phytases with improved properties such as temperature stability, pH stability and substrate specificity, for use in pharmaceuticals and compound foods and feeds -

PT

PS Claim 10; Fig 7a-c; 240pp; English.

XX

CC The present invention describes improved phytases, preferably with increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and end level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence encodes a phytase sequence from the present invention.

XX

SO Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 1404; DB 21; Length 1404;

Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggggctgttcgtgctactgtccattgccacctgttcggttccacatccgatacc 60

DB 1 atggggctgttcgtgctactgtccattgccacctgttcggttccacatccgatacc 60

QY 61 gcttgggtctcctggtgaatttcactctgtgacactgttgacggtgttaccatgt 120

DB 61 gcttgggtctcctggtgaatttcactctgtgacactgttgacggtgttaccatgt 120

QY 121 ttccagaaatttcactctgtggtggtacactctccactctctctttggcagacgaa 180

DB 121 ttccagaaatttcactctgtggtggtacactctccactctctctttggcagacgaa 180

QY 181 tctgctattctccagacgttccagacgactgtagagttacttctgtaagtgtct 240

DB 181 tctgctattctccagacgttccagacgactgtagagttacttctgtaagtgtct 240

QY 241 agacaggtgtctagataccacactctctcgtctaaaggttactctctgttgatgaa 300

DB 241 agacaggtgtctagataccacactctctcgtctaaaggttactctctgttgatgaa 300

QY 301 gctattccaaagacgctactgctttccaaaggttaagtagctttcttgagacttacac 360

DB 301 gctattccaaagacgctactgctttccaaaggttaagtagctttcttgagacttacac 360

QY 361 tacatttgggtctcagactgtacatccattcgttgaaacaaatggttaactctggt 420

DB 361 tacatttgggtctcagactgtacatccattcgttgaaacaaatggttaactctggt 420

QY 421 attaatctctacagaagatcacaggctttggtctagaagattgtccattcattagact 480

DB 421 attaatctctacagaagatcacaggctttggtctagaagattgtccattcattagact 480

QY 481 tctggttctgacagagtattgtctctgctgaaagttcattgaaggtttccaactctct 540

DB 481 tctggttctgacagagtattgtctctgctgaaagttcattgaaggtttccaactctct 540

QY 541 aagttggtgacccaggttctcaacacacaaagcttctccagttattacagtgatcatt 600

DB 541 aagttggtgacccaggttctcaacacacaaagcttctccagttattacagtgatcatt 600

QY 601 ccagaaggtatccggtttacaaacacactttggaccacggttactgttcttccgaagac 660

DB 601 ccagaaggtatccggtttacaaacacactttggaccacggttactgttcttccgaagac 660

QY 661 tctgaattagtagtgcagcgttgaaacttaacttcaactgcttcttccagctattaga 720

DB 661 tctgaattagtagtgcagcgttgaaacttaacttcaactgcttcttccagctattaga 720

QY 721 gctagattggaagctgacttgcaggtgttactttgactgacgaaagacgttcttacttg 780

DB 721 gctagattggaagctgacttgcaggtgttactttgactgacgaaagacgttcttacttg 780

QY 781 atggacatgttccattcgcactgttcgtagaacttctgacgctactgaattgtctcca 840

DB 781 atggacatgttccattcgcactgttcgtagaacttctgacgctactgaattgtctcca 840

QY 841 tctgtgcttcttcaactcacacaaatgataccaaacttaacttgcagctactgcaagctgggt 900

DB 841 tctgtgcttcttcaactcacacaaatgataccaaacttaacttgcagctactgcaagctgggt 900

QY 901 aagtaactcgggttgcggtgtgtaacccattgggttccagctcaaggtgtgttctcgt 960

DB 901 aagtaactcgggttgcggtgtgtaacccattgggttccagctcaaggtgtgttctcgt 960

QY 961 aacgaattgattgctagattgactcactctccagttccagttccagcttacttactaacac 1020

DB 961 aacgaattgattgctagattgactcactctccagttccagttccagcttacttactaacac 1020

QY 1021 actttggactctaaacccagctacttccattgaacgtacttcttgcagctacttctcgt 1080

DB 1021 actttggactctaaacccagctacttccattgaacgtacttcttgcagctacttctcgt 1080

QY 1081 caccgacacactatgatattcttcttctggtttgttgcacacggttaccagcca 1140

DB 1081 caccgacacactatgatattcttcttctggtttgttgcacacggttaccagcca 1140

QY 1141 ttgtctactactctctgtgaactctattgaagaaactgacggttactctcttctggact 1200

DB 1141 ttgtctactactctctgtgaactctattgaagaaactgacggttactctcttctggact 1200

QY 1201 gttccattcgtctagagcttaccgttgaatgacaaatgacaaactgcaagctgaaagaccca 1260

DB 1201 gttccattcgtctagagcttaccgttgaatgacaaatgacaaactgcaagctgaaagaccca 1260

QY 1261 ttgggttagagtttgggttaacgacagagtttccattgcacggttgcgtgttgacaag 1320

DB 1261 ttgggttagagtttgggttaacgacagagtttccattgcacggttgcgtgttgacaag 1320

QY 1321 ttgggttagagtttgggttaacgacagagtttccattgcacggttgcgtgttgacaag 1380

DB 1321 ttgggttagagtttgggttaacgacagagtttccattgcacggttgcgtgttgacaag 1380

QY 1381 aactgggctgaattgttctgcttaa 1404

DB 1381 aactgggctgaattgttctgcttaa 1404

RESULT 3

AAZ59715

ID AAZ59715 standard; DNA; 1404 BP.

XX AAZ59715;

AC AAZ59715;

XX AAZ59715;

DT 19-APR-2000 (first entry)

XX DNA

DE DNA encoding a mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;

KW thermostable; animal feed; monogastric animal; phytate phosphorus;

KW phosphate availability; consensus; mutant; ds.

XX Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS	Aspergillus niger var. awamori.	
OS	Aspergillus niger str. NRRL3135.	
OS	Aspergillus fumigatus ATCC13073.	
OS	Aspergillus fumigatus ATCC32722.	
OS	Aspergillus fumigatus ATCC58128.	
OS	Aspergillus fumigatus ATCC26906.	
OS	Aspergillus fumigatus ATCC32239.	
OS	Emericella nidulans.	
OS	Talaromyces thermophilus ATCC20186.	
OS	Myceliophthora thermophila.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
FT	CDS	1..1404
FT		/*tag= a
FT		/product= "Phytase-1-thermo[8]-Q50T-K91A"
XX	EP969089-AL.	
XX	05-JAN-2000.	
XX	23-JUN-1999;	99EP-0111949.
XX	29-JUN-1998;	98EP-0111960.
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX	Brugger R, Lehmann M, Wyss M;	
XX	WPI: 2000-099429/09.	
DR	P-PSDB; AAY69568.	
DR	New stabilized enzyme formulation, useful for feed compositions for monogastric animals -	
PT	Example 5; Fig 19; 101pp; English.	
PS	The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate CC phosphohydrolase) and one or more stabilising agents including CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic, CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate. CC The stabilised phytase formulation is used in a method for preparing a CC feed composition for monogastric animals (e.g., pigs, poultry) and CC provides a monogastric animal with its dietary requirements of CC phosphorus. Although a large amount of phosphate is present in animal CC feed in the form of phytate phosphorus, monogastric animals are unable CC to utilise this form of phosphate, resulting in the addition of extra CC phosphate to the feed of such animals. Phytase enhances the nutritional CC value of plant material without the need for adding additional phosphate CC to the feed. The level of phosphate pollution in the environment is CC reduced by adding phytase to animal feed, as the animal can make use of CC the inorganic phosphate liberated from phytate phosphorus using the CC enzyme. The phytase formulation of the invention has an improved CC thermostability and can therefore remain stable during long-term storage CC and can withstand feed processing methods such as extrusion, expansion CC and pelleting. The present sequence represents DNA encoding a mutant CC phytase-1 consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a CC temperature optimum and melting point 7 degrees Celsius higher than that CC of phytase-1 (AAY69558).	
XX	Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;	
SQ		
Query Match 100.0%; Score 1404; DB 21; Length 1404;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 atggcggtgctcggtgctactgtccattgccaccttggttcggttcacatccggtacc 60	
Db	1 atggcggtgctcggtgctactgtccattgccaccttggttcggttcacatccggtacc 60	

Increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. thermostability, pH stability, pH profile, temperature profile, temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence encodes a phytase sequence from the present invention.

Sequence 1404 BP: 330 A: 320 C: 302 G: 452 T: 0 other;

Query Match	98.6%	Score 1384.8;	DB 21;	Length 1404;
Best Local Similarity	99.1%	Pred. No. 0;		
Matches 1392;	Conservative	0; Mismatches	12;	Indels 0;

QY	1	atggggcgttgctgcgtcactgtccattgccaccctgttcgattccacatccggtacc	60
Db	1	atggggcgtttctgcgtcactgtccattgccaccctgttcgattccacatccggtacc	60
QY	61	gccttgggtctcgtgtgaattcactctgtgacactgttgacggtgggttaccaatgt	120
Db	61	gccttgggtctcgtgtgaattcactctgtgacactgttgacggtgggttaccaatgt	120
QY	121	ttcccagaaattctcacttctgggttactactctccatactctcttttggcagacgaa	180
Db	121	ttcccagaaattctcacttctgggttactactctccatactctcttttggcagacgaa	180
QY	181	tctgtatttccagacggtccagacgacttagaagttaacttcttcgttcaagttttgtct	240
Db	181	tctgtatttccagacggtcccaaaaggacttagaagttaacttcttcgttcaagttttgtct	240
QY	241	agacacggtgctagatacccaactctcttcgcttaagcttactctgtcttgcattgtaa	300
Db	241	agacacggtgctagatacccaactctcttaagcttaagcttactctgtcttgcattgtaa	300
QY	301	gctattcaaaagaacgtactgtcttcaaggtaagtaogctttcttgaagacittacaac	360
Db	301	gctattcaaaagaacgtactgttcaaggtaagtaogctttcttgaagacittacaac	360
QY	361	tacactttgggtgctgacgacgtgactcccatcgggtgaaacccaatggttaactctggt	420
Db	361	tacactttgggtgctgacgacgtgactcccatcgggtgaaacccaatggttaactctggt	420
QY	421	attaagttctacagaagatacaaggctttggcttagaagaagtgttccattccattagagct	480
Db	421	attaagttctacagaagatacaaggctttggcttagaagaagtgttccattccattagagct	480
QY	481	tctggttctgcagagattattgtctctgtgaaaaagttcaatgaagggttccaatctgct	540
Db	481	tctggttctgcagagattattgtctctgtgaaaaagttcaatgaagggttccaatctgct	540
QY	541	aagttgctgcacccaggttctcaaccacaccaaagcttctccagttattaaacytgatcatt	600
Db	541	aagttgctgcacccaggttctcaaccacaccaaagcttctccagttattaaacytgatcatt	600
QY	601	ccagaaggatccgggttacaaacacactttggacaacogtgtaactgtactgcttctcaagac	660
Db	601	ccagaaggatccgggttacaaacacactttggaccatggtcttctgtactgcttctcaagac	660
QY	661	tcTgaattagggtgacgacggttgaagctaaactcaactcttctgtccctcagctattaga	720
Db	661	tctacactcaggttgaacgaggttgaagctaaactcaactcttctgtccctcagctattaga	720
QY	721	gctagattggaagctgacttgcacggtgttactttgactgacgaagaacggttttacttctg	780
Db	721	gctagattggaagctgacttgcacggtgttactttgactgacgaagaacggttttacttctg	780
QY	781	atggacatgtgtccattccgacactgtgctagaactctctgaogctactgaaattgtgtccca	840

[illegible]

05-DEC-2000	(first entry)
Consensus phytase 3 thermo 11 Q50T polynucleotide SEQ ID NO:90.	
<p> phytase; mutant; thermostability; mutation; mutagenesis; pH stability; temperature stability; pH profile; temperature profile; reaction rate; specific activity; substrate specificity; substrate cleavage pattern; substrate binding; position specificity; phytate degradation rate; food; feed; phytate; manure; ds. </p>	
Synthetic.	
WO2000043503-A1.	
27-JUL-2000.	
21-JAN-2000; 2000WO-DK00025.	
22-JAN-1999; 99DK-0000092.	
21-SEP-1999; 99DK-0001340.	
(NOVO) NOVO NORDISK AS.	
Lehmann M;	
WPI; 2000-491161/43.	
P-PSDB; AAB20531.	
<p> Novel phytases with improved properties such as temperature stability, pH stability and substrate specificity, for use in pharmaceuticals and compound foods and feeds - </p>	
Disclosure; Fig 22a-c; 240pp; English.	
The present invention describes improved phytases, preferably with	

Qy	481	tctggtcttgacagagttattgcttctgctgaaagttcattgaaggtttccaactcgt	54
Db	492	tctggtcttgacagagttattgcttctgctgaaagttcattgaaggtttccaactcgt	551
Qy	541	aagttggtgaccaggtttctcaaccacacaagcttctccagttattaacgtgatcatt	600
Db	552	aagttggtgaccaggtttctcaaccacacaagcttctccagttattaacgtgatcatt	611
Qy	601	ccagaagatccggtttacaacaacactttgacacacggttacttctactccttcaaac	660
Db	612	ccagaagatccggtttacaacaacactttgacacacggttacttctactccttcaaac	671
Qy	661	tctgaattaggtgacgacgttgaagctaaacttcaactgcttctgctccagctattaga	720
Db	672	tctgaattggtgacgacgttgaagctaaacttcaactgcttctgctccagctattaga	731
Qy	721	gctagattgaagctgacttgcaggtgttactttgactgacgaagcgttttactctg	780
Db	732	gctagattggaagctgacttgcaggtgttactttgactgacgaagcgttttactctg	791
Qy	781	atgacatgtgtccacttcgacacttcgctagaacttcagcaacttcagcgtactgaattgctcca	840
Db	792	atgacatgtgtccacttcgcaaaactgtgtctagaacttcgacgtactgaattgctcca	851
Qy	841	tctctgcttcttcaactcacgacgaatggatccaactacgacttctgcaagcgttggt	900
Db	852	tctctgcttcttcaactcacgacgaatggatccaactacgacttctgcaagcgttggt	911
Qy	901	aagttactacggttacggtgctgaaccatttgggtccagctcaaggtgttggttgcct	960
Db	912	aagttactacggttacggtgctgaaccatttgggtccagctcaaggtgttggttgcct	971
Qy	961	aacgaattgattgctagattgactcactctccagttcaagacacacacttctactaacac	1020
Db	972	aacgaattgattgctagattgactcactctccagttcaagacacacacttctactaacac	1031
Qy	1021	acttggactcaccagactacttcccattgaacgtactttgacgtgacttctct	1080
Db	1032	acttggactcaccagactacttcccattgaacgtactttgacgtgacttctct	1091
Qy	1081	cacgaacaactatgatattctatttctcgtgttgggtttgtacaacggtaccaaagcca	1140
Db	1092	cacgaacaactatgatattctatttctcgtgttgggtttgtacaacggtaccaaagcca	1151
Qy	1141	tgtctactactctgttgaactctattgaagaaactgacgttactctgcttctgact	1200
Db	1152	tgtctactactctgttgaactctattgaagaaactgacgttactctgcttctgact	1211
Qy	1201	gttccattcgtcgtgacttacgtttgaattgatgcaatgtcaagctgaaaaggaacca	1260
Db	1212	gttccattcgtcgtgacttacgtttgaattgatgcaatgtcaagctgaaaaggaacca	1271
Qy	1261	tgggttagagtttgggttaacgacagagttgtccattgacggttgtgttgacaag	1320
Db	1272	tgggttagagtttgggttaacgacagagttgtccattgacggttgtgttgacaag	1331
Qy	1321	tgggttagatgtaagagagacacttcgttgaaggttcttcttcgctaactcgttggt	1380
Db	1332	tgggttagatgtaagagagacacttcgttgaaggttcttcttcgctaactcgttggt	1391
Qy	1381	aactgggctgaatgtttcgtctaa	1404
Db	1392	aactgggctgaatgtttcgtctaa	1415

RESULT 8
AAA73231
ID AAA73231 standard; DNA; 1426 BP.
XX
XX AAA73231;
XX AC
DT 05-DEC-2000 (first entry)

XX	Consensus phytase polynucleotide sequence SEQ ID NO:15.
DE	
XX	Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW	temperature stability; pH profile; temperature profile; reaction rate;
KW	specific activity; substrate specificity; substrate cleavage pattern;
KW	substrate binding; position specificity; phytate degradation rate;
KW	food; feed; phytate; manure; ds.
XX	
OS	Synthetic.
XX	
PN	WO2000043503-A1.
XX	
PD	27-JUL-2000.
XX	
PF	21-JAN-2000; 2000WO-DK00025.
XX	
PR	22-JAN-1999; 99DK-0000092.
PR	21-SEP-1999; 99DR-0001340.
XX	
PA	(NOVO) NOVO NORDISK AS.
XX	
PI	Lehmann M;
XX	
DR	WPI; 2000-491161/43.
DR	P-PSDB; AAB20515.
XX	
PT	Novel phytases with improved properties such as temperature stability;
PT	pH stability and substrate specificity, for use in pharmaceuticals and
PT	compound foods and feeds -
XX	
PS	Example 9; Fig 2a-c; 240pp; English.
XX	
CC	The present invention describes improved phytases, preferably with
CC	increased thermostability, and methods for producing them. The methods
CC	can be used for producing phytases with improved properties e.g. file,
CC	temperature stability, pH stability, pH profile, temperature profile,
CC	specific activity, substrate specificity, substrate cleavage pattern,
CC	substrate binding, position specificity, the velocity and level of
CC	release of phosphate from corn, reaction rate, phytate degradation rate
CC	and end level of released phosphate. The phytases can be used to produce
CC	pharmaceutical compositions or compound food or feeds. The feed can be
CC	used to reduce levels of phytate in animal manure, by converting it
CC	into lower inositol phosphates and/or inositol and inorganic phosphate
CC	The present sequence encodes a phytase sequence from the present
CC	invention.

Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;

Query Match	96.94%	Score 1360.8	DB 21	Length 1426
Best Local Similarity	98.1%	Pred. No. 0		
Matches 1377	Conservative 0	Mismatches 27	Indels 0	Gaps 0
QY	1	atggggcgtgttcgtgctactgtccattgccacattgttcgcgttccacatcccggtacc	60	
Db	12	atgggcgtgttcgtgctactgtccattgccacattgttcgcgttccacatcccggtacc	71	
QY	61	gccttgggtctcgttggaatttcacactctctgtgacactgttgacgggtgggtaccaatgt	120	
Db	72	gccttgggtctcgttggaatttcacactctctgtgacactgttgacgggtgggtaccaatgt	131	
QY	121	ttccacagaaattctcacttgggtgacctactctccatactctctcttggcagacgaa	180	
Db	132	ttccacagaaattctcacttgggtgacctactctccatactctctcttggcagacgaa	191	
QY	181	tctgctatttccacagcttccacagacgtgtgaggttacttcttcgaagttttgtct	240	
Db	192	tctgctatttccacagcttccacagacgtgtgaggttacttcttcgaagttttgtct	251	
QY	241	agacacggtctagataccacaaacttcttcgtctctaaagcttactctgctttgattgaa	300	
Db	252	agacacggtctagataccacaaacttcttcgaagcttactctgctttgattgaa	311	

PR	22-JAN-1999;	99DK-0000091.	
PR	22-JAN-1999;	99DK-0000093.	
XX	PA	(NOVO) NOVO-NORDISK AS.	
XX	PI	Petersen S;	
XX	XX	WPI; 1999-591030/50.	
DR	DR	P-PSDB; AAY43170.	
XX	PT	Preparing animal feed using a thermostable phytase	
XX	XX	Example 3; Fig 10; 71pp; English.	
XX	XX	This sequence encodes the consensus phytase-10-thermo(3)-Q50T-K91A.	
CC	CC	The invention relates to a process for preparing animal feed by	
CC	CC	agglomerating feed ingredients with a thermostable phytase, which is	
CC	CC	added before or during agglomeration. The thermostable phytase is useful	
CC	CC	for expression in transgenic plants. These plants are useful in the	
CC	CC	preparation of animal feed itself. The thermostable phytase allows animal	
CC	CC	feed to be produced more efficiently, in addition to improved	
CC	CC	phytase-expressing transgenic plants. These plants provide a feed	
CC	CC	ingredient and a feed additive (phytase) simultaneously.	
XX	XX	Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;	
SQ			
	Query Match	94.5%; Score 1327.2; DB 20; Length 1404;	
	Best Local Similarity	96.6%; Pred. No. 0;	
	Matches 1356; Conservative	0; Mismatches 48; Indels 0; Gaps 0;	
QY	1	atggcggtgttcgtgctactgtccattgccacctgttcgggtccacatccgggtacc	60
DB	1	atggcggtgttcgtgctactgtccattgccacctgttcgggtccacatccgggtacc	60
QY	61	gcttgggtccctcgtggttaattctcactcttctgacactgttgacggttggttaccatgt	120
DB	61	gcttgggtccctcgtggttaattctcactcttctgacactgttgacggttggttaccatgt	120
QY	121	ttcccaagaattcttcactgttgggtgaactactctcactctcactcttcttggcgagacgaa	180
DB	121	ttcccaagaattcttcactgttgggtgaactactctcactctcactcttcttgggtgagacaa	180
QY	181	ttctgctattctccagacgttccagacgactgtagattactttcgttcaagattttgtct	240
DB	181	ttctgctattctccagacgttccagaggtgttagattactttcgttcaagattttgtct	240
QY	241	agacacggtgtctagataccacacttcttctcgtctgaaggttactctgtttgattgaa	300
DB	241	agacacggtgtctagataccacacttcttctcgtctgaaggttactctgtttgattgaa	300
QY	301	gtattcaaaaagaacgtactgttttcaagggttaagtacgctttcttgaagacttacac	360
DB	301	gtattcaaaaagaacgtactgttttcaagggttaagtacgctttcttgaagacttacac	360
QY	361	tacacttgggtgtcgtacgactgtactcattcgttgaaacacaaatggttaactctgt	420
DB	361	tacacttgggtgtcgtacgactgtactcattcgttgaaacacaaatggttaactctgt	420
QY	421	attaagttctacagaagatacaagggtttggtgtagaagattgttcattcattagact	480
DB	421	attaagttctacagaagatacaagggtttggtgtagaagattgttcattcattagact	480
QY	481	ttcgtttctgacagattattgtctcgtcgtgaaagtctcattgaaagtttccaaatctgt	540
DB	481	ttcgtttctgacagagctattgtctcgtcgtgaaagtctcattgaaagtttccaaatctgt	540
QY	541	aagtgtgctgaccaggttctcaccacacaaagcttctccagttatttaacgtatcat	600
DB	541	aagtgtgctgaccaggttctcaccacacaaagcttctccagttatttaacgtatcat	600
QY	601	ccagaagagatccggtttacaacacactttggaccacggtactgttactgttcgaagac	660
DB	601	ccagaagaggtgctggtttacaacacactttggaccacggtttgttactgttctcgaagaa	660
QY	661	tctgaattagtgacgacgttgaagcttaactcactcactgcttcttctcgtcccgctattaga	720
DB	661	tctgaattgggtgacgacgttgaagcttaactcactcactgcttcttctcgtcccgctattaga	720
QY	721	gctgattggaagctgacttgcagggtgttacttgaactgacgagacgttgtttactgt	780
DB	721	gctgattggaagctcacttgcagggtgttacttgaactgacgagacgttgtttactgt	780
QY	781	atggacatgttccattcgacactgttctgtagaacttctgacgtactcaactcgtctcca	840
DB	781	atggacatgttccattcgacactgttctgtagaacttctgacgtactcaactcgtctcca	840
QY	841	tctgtgcttcttactcactcagcagaaatggtccaatcagcactacttgcgaagcttgggt	900
DB	841	tctgtgacttcttactcactcagcagaaatggtccaatcagcactacttgcgaagcttgggt	900
QY	901	aagtactacggttacggtgctggttaacccattgggtccagctcaagggtgttgggttcgt	960
DB	901	aagtactacggttacggtgctggttaacccattgggtccagctcaagggtgttgggttcgt	960
QY	961	aacgaattgattgctagattgactcactcctccagttccagaccacacttctactaaccac	1020
DB	961	aacgaattgattgctagattgactcactcctccagttccagaccacacttctactaaccac	1020
QY	1021	actttgactctaacccagctacttcccatgaacgctactttgtacgctgactctct	1080
DB	1021	actttgactctaacccagctacttcccatgaacgctactttgtacgctgactctct	1080
QY	1081	cagcacaacactatgatatctattttctcgttttgggtttgtacacggttaccagcca	1140
DB	1081	cagcacaacactatggtttctattttctcgttttgggtttgtacacggttaccagcca	1140
QY	1141	ttgtctactacttctgttaaatctattgaagaaactgacggttactctctcttggact	1200
DB	1141	ttgtctactacttctgttgaatctattgaagaaactgacggttactctctcttggact	1200
QY	1201	gttcattcgtcgtcgtacgcttacttgaatgatcaatgtcaagctgaagaaagaccca	1260
DB	1201	gttcattcgtcgtcgtacgcttacttgaatgatcaatgtcaagctgaagaaagaccca	1260
QY	1261	ttggttagagttttgttaacgacagagttgttccattgacggttctgtctgacaag	1320
DB	1261	ttggttagagttttgttgaacgacagagttgttccattgacggttctgtctgacaag	1320
QY	1321	ttggttagatgtagagacgacttctcgttgaaggtttgtcttctcgtcgtatcgtggt	1380
DB	1321	ttggttagatgtagagacgacttctcgttgaaggtttgtcttctcgtcgtatcgtggt	1380
QY	1381	aactgggtggaattttcgtctaa	1404
DB	1381	aactgggaagaattttcgtctaa	1404
XX	AAAY73234;		
XX	XX	05-DEC-2000 (first entry)	
XX	XX	Consensus phytase 10 thermo 3 q50t, k91a polynucleotide SEQ ID NO:30.	
XX	XX	Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;	
XX	XX	temperature stability; pH profile; temperature profile; reaction rate;	
XX	XX	specific activity; substrate specificity; substrate cleavage pattern;	
XX	XX	substrate binding; position specificity; phytate degradation rate;	
XX	XX	food; feed; phytate; manure; ds.	
XX	XX	Synthetic.	

RESULT 11
AAAY73234
ID AAA73234 standard; DNA; 1404 BP.

PN WO200043503-A1.
 XX 27-JUL-2000.
 PD 21-JAN-2000; 2000WO-DK00025.
 PF 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PA Lehmann M;
 PI WPI; 2000-491161/43.
 DR P-PSDB; AAB20527.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX Claim 10; Fig 8a-c; 240pp; English.
 PS The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence encodes a phytase sequence from the present
 CC invention.
 XX Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;
 SQ

Query Match 94.5%; Score 1327.2; DB 21; Length 1404;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1356; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 atggcggtgttcgctgactgtccattgcacactgttgctggttccacatccggtacc 60
 DB 1 atggcggtgttcgctgactgtccattgcacactgttgctggttccacatccggtacc 60
 QY 61 gcttgggttcctggtgtaattctcactctgtgacactgttgacactgttgacactgtg 120
 DB 61 gcttgggttcctggtgtaattctcactctgtgacactgttgacactgttgacactgtg 120
 QY 121 ttccagaaattctcactgttggtggtgactctcactctcactctcactctcactctcact 180
 DB 121 ttccagaaattctcactgttggtggtgactctcactctcactctcactctcactctcact 180
 QY 181 tctgctattcttcacagctgtccagacgactgtagactgttacttctcactctcactctcact 240
 DB 181 tctgctattcttcacagctgtccagaggtgtcagaggttacttctcactctcactctcact 240
 QY 241 agacacggtgtctagatgccaaactctctcgtgtcgtctcgtctcgtctcgtctcgtctgaa 300
 DB 241 agacacggtgtctagatgccaaactctctcgtgtcgtctcgtctcgtctcgtctcgtctgaa 300
 QY 301 gctattcaaaagacgactactgttccaggttaagtaagcgtcttcttgaagacttacacac 360
 DB 301 gctattcaaaagacgactactgttccaggttaagtaagcgtcttcttgaagacttacacac 360
 QY 361 tacacttgggtgctgacgactgtactcattcgtggtgaaacacaaatggttaactctggt 420
 DB 361 tacacttgggtgctgacgactgtactcattcgtggtgaaacacaaatggttaactctggt 420
 QY 421 attaatctacagagatcacaggcttggctgtagaagattgttccattcattagagct 480

Db 421 attaatctacagagatcacaggcttggctgtagaagattgttccattcattagagct 480
 QY 481 tctggttctgacagagatttctctgctgaaagtctcattgaaaggtttccaatctgct 540
 Db 481 tctggttctgacagagatttctctgctgaaagtctcattgaaaggtttccaatctgct 540
 QY 541 aagtgtgctgacccaggttctcaccacacacacacacacacacacacacacacacacacac 600
 Db 541 aagtgtgctgacccaggttctcaccacacacacacacacacacacacacacacacacacac 600
 QY 601 ccagaaaggtcgggttac 660
 Db 601 ccagaaaggtcgggttac 660
 QY 661 tctgaattagggtgacgacgttgagcttaacttcaacttcaacttcaacttcaacttcaacttca 720
 Db 661 tctgaattagggtgacgacgttgagcttaacttcaacttcaacttcaacttcaacttcaacttca 720
 QY 721 gctagattggaagctgacttgcaggtgttaacttgaactgacgaaagcgttcttacttctg 780
 Db 721 gctagattggaagctgacttgcaggtgttaacttgaactgacgaaagcgttcttacttctg 780
 QY 781 atggacatgttccattcgcacactgtcgtgtagaacttctgacgactgacgactgacgactgac 840
 Db 781 atggacatgttccattcgcacactgtcgtgtagaacttctgacgactgacgactgacgactgac 840
 QY 841 tctgtgcttcttcaactcagcagcaatggtatccaatcagcactacttgcacacacacacacac 900
 Db 841 tctgtgacttcttcaactcagcagcaatggtatccaatcagcactacttgcacacacacacacac 900
 QY 901 aagtactcgggttgcgtgctgtaaccacacacacacacacacacacacacacacacacacacac 960
 Db 901 aagtactcgggttgcgtgctgtaaccacacacacacacacacacacacacacacacacacacac 960
 QY 961 aacgaattgattgctagattgactcactctcactctcactctcactctcactctcactctcactc 1020
 Db 961 aacgaattgattgctagattgactcactctcactctcactctcactctcactctcactctcactc 1020
 QY 1021 accttggactctaac 1080
 Db 1021 accttggactctaac 1080
 QY 1081 cagcacaac 1140
 Db 1081 cagcacaac 1140
 QY 1141 ttgtctactactcttctgtaacttattgaaagaaacacacacacacacacacacacacacac 1200
 Db 1141 ttgtctactactcttctgtaacttattgaaagaaacacacacacacacacacacacacacac 1200
 QY 1201 gttccattcgtcgtgactgacttgcgttgaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1260
 Db 1201 gttccattcgtcgtgactgacttgcgttgaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1260
 QY 1261 ttggttagagtttgggttaacgacagagttgttccattgacacacacacacacacacacacac 1320
 Db 1261 ttggttagagtttgggttaacgacagagttgttccattgacacacacacacacacacacacac 1320
 QY 1321 ttgggttagatgtaagagagacgacttgcgttgaaggtttgttcttgcgtagatctggttgg 1380
 Db 1321 ttgggttagatgtaagagagacgacttgcgttgaaggtttgttcttgcgtagatctggttgg 1380
 QY 1381 aactggggtcgaatttctcgtctaa 1404
 Db 1381 aactggggaagaatttctcgtctaa 1404

RESULT 12
 AAZ59716
 ID AAZ59716 standard; DNA; 1404 BP.
 XX AC
 XX AC
 XX AC

Db 901 aagtactacggttagctggttaacccattgggtccagctcaagggtgtgtggtttcgtt 960
 QY 961 aacgaattgattgctagattgaactcaactctccagttcaagaccacactcttactaacac 1020
 Db 961 aacgaattgattgctagattgaactcaactctccagttcaagaccacactcttactaacac 1020
 QY 1021 actttggactctaacccagctacttccattgaacgctactttgtacgctgacttctct 1080
 Db 1021 actttggactctaacccagctacttccattgaacgctactttgtacgctgacttctct 1080
 QY 1081 cagcaaacactatgatactattttctctcgtttgttggtttgtacaaacggtaccagcca 1140
 Db 1081 cagcaaacactatggtttctattttctctcgtttgttggtttgtacaaacggtaccagcca 1140
 QY 1141 ttgtctactactctgttggaactattgaagaaactgacggttactctgtctcttgact 1200
 Db 1141 ttgtctactactctgttggaactattgaagaaactgacggttactctgtctcttgact 1200
 QY 1201 gtccattcgcgttagagctacgttgaaatgatgcaatgtcaagctgaaaggaacca 1260
 Db 1201 gtccattcgcgttagagctacgttgaaatgatgcaatgtgaagctgaaaggaacca 1260
 QY 1261 ttggttagattttgttgaacgacagattgttccattgaacggttgtgtgttgacaag 1320
 Db 1261 ttggttagattttgttgaacgacagattgttccattgaacggttgtgtgttgacaag 1320
 QY 1321 ttggttagattgaagagacgacttcgttgaaagttgttcttctcgttagatctgtgtgt 1380
 Db 1321 ttggttagattgaagagacgacttcgttgaaagttgttcttctcgttagatctgtgtgt 1380
 QY 1381 aactgggctgaattttcgttaa 1404
 Db 1381 aactgggaagaattgttcgcttaa 1404

RESULT 13

AAA73293

ID AAA73293 standard; DNA; 1404 BP.

XX AC AAA73293;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 10 thermo 5 Q50T K91A polynucleotide SEQ ID NO:96.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 XX food; feed; phytate; manure; ds.

XX OS Synthetic.

XX PN WO200043503-A1.

XX XX 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX XX WPI; 2000-491161/43.

XX DR P-PSDB; AAB20534.

XX PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX XX

PS Disclosure; Fig 25a-c; 240pp; English.

XX CC The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence encodes a phytase sequence from the present
 CC invention.

XX SQ Sequence 1404 BP; 329 A; 312 C; 304 G; 459 T; 0 other;

Query Match 94.3%; Score 1324; DB 21; Length 1404;

Best Local Similarity 96.4%; Pred. No. 0;

Matches 1354; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 atggcggtgttcgtcgtactgtccattgccactgttgcgttccacatcccggtacc 60
 Db 1 atggcggtgttcgtcgtactgtccattgccactgttgcgttccacatcccggtacc 60
 QY 61 gcoctgggttcctcgttgtaattctcactcttgcacactgttgacggtgttaccatgt 120
 Db 61 gcoctgggttcctcgttgtaattctcactcttgcacactgttgacggtgttaccatgt 120
 QY 121 ttcccagaattttcactgttgggtaccactactctccatactcttctttggcagacgaa 180
 Db 121 ttcccagaattttcactgttgggtaccactactctccatactcttctttggcagacgaa 180
 QY 181 tctgtattttccagacgcttccagacgacttagattacttcttcgttcaagtttgtct 240
 Db 181 tctgtattttccagacgcttccagacgacttagattacttcttcgttcaagtttgtct 240
 QY 241 agacacggtgctagatatacccaacttcttcgtcgttaaggcttactctgtttgattgaa 300
 Db 241 agacacggtgctagatatacccaacttcttcgtcgttaaggcttactctgtttgattgaa 300
 QY 301 gctattccaaaagacgctactgtttccaaaggtaagtaagcgtttcttgaagacttacaac 360
 Db 301 gctattccaaaagacgctactgtttccaaaggtaagtaagcgtttcttgaagacttacaac 360
 QY 361 tacactttgggtgctgacgacttgactccattcgttgaaacacaaatggttaactctgt 420
 Db 361 tacactttgggtgctgacgacttgactccattcgttgaaacacaaatggttaactctgt 420
 QY 421 attaatgtctacagaagatacaaaaggctttggttagaagaattgttccattcattagact 480
 Db 421 attaatgtctacagaagatacaaaaggctttggttagaagaattgttccattcattagact 480
 QY 481 tctgttctcagacaggttatgttcttcgtcaaaaagttcattgaaggtttccactctgt 540
 Db 481 tctgttctcagacaggttatgttcttcgtcaaaaagttcattgaaggtttccactctgt 540
 QY 541 aagttggctgacccaggttctcaaccacacaaagcttctccagttattaaactgtatcatt 600
 Db 541 aagttggctgacccaggttctcaaccacacaaagcttctccagttattaaactgtatcatt 600
 QY 601 ccagaaggtacggtttacacaaacactttggaccacggttactgtactgtttcgaagac 660
 Db 601 ccagaaggtacggtttacacaaacactttggaccacggttactgtactgtttcgaagac 660
 QY 661 tctgaattaggtgacgacgtttgaagctaaactcactcttctgttgcgtccagctattaga 720
 Db 661 tctaccctaggtgacgacgtttgaagctaaactcactcttctgttgcgtccagctattaga 720
 QY 721 gctagattggaagctgacttgcaggtgttactttgactgacgaagacgttgttactgtt 780
 Db 721 gctagattggaagctgacttgcaggtgttactttgactgacgaagacgttgttactgtt 780

Job time: 5033 sec

Db 372 tacactttgggtgtgtgacgaacttgactccattcgttgaaacacaaatggttaactctggt 431
Qy 421 attaagttctacagaagatacaagcttttggctagaaagattttccattcattagagct 480
|||||
Db 432 attaagttctacagaagatacaagcttttggctagaaagattttccattcattagagct 491
|||||
Qy 481 tctggttctgacagaggttattgtctctctgctgaaagttcattgaaggtttccaatctgct 540
|||||
Db 492 tctggttctgacagaggttattgtctctctgctgaaagttcattgaaggtttccaatctgct 551
|||||
Qy 541 aagttggtgacccaggttctcaacacacacagcttctccagttattaaacgtgatcatt 600
|||||
Db 552 aagttggtgacccaggttctcaacacacacagcttctccagttattaaacgttatatt 611
|||||
Qy 601 ccagaaggtacccggtttacaacacacactttggaccacggttactgttactgttccgaagac 660
|||||
Db 612 ccagaaggtgctggtttacaacacacactttggaccacggttactgttactgttccgaagaa 671
|||||
Qy 661 tctgaattaggtgacagaggttgaagctaaacttcaactcactgtttgttgcgtccagctattaga 720
|||||
Db 672 tctgaattggtgacagaggttgaagctaaacttcaactcactgtttgttgcgtccacattataga 731
|||||
Qy 721 gctagattggaagctgacttgcaggtgttactttgactgacgaagacgttgttacttg 780
|||||
Db 732 gctagattggaagctcacttgcaggtgttactttgactgacgaagacgttgttacttg 791
|||||
Qy 781 atggacatgtgtccattccacactgtcgtcgtagaaacttctgacgtactgaaattgtctcca 840
|||||
Db 792 atggacatgtgtccattccacactgtcgtcgtagaaacttctgacgtactgaaattgtctcca 851
|||||
Qy 841 tctctgtcttcttcaactcacacgaatggtatccaatcagactacttgcagaagcttgggt 900
|||||
Db 852 tctctgtcactgttcaactcacacgaatggtatccaatcagactacttgcagaagcttgggt 911
|||||
Qy 901 aagtaactcaggttacggtgtgtgtacccattgggtccagctcgaaggtgtgtgttgcgt 960
|||||
Db 912 aagtaactcaggttacggtgtgtgtacccattgggtccagctcgaaggtgtgtgttgcgt 971
|||||
Qy 961 aacgaattgattgctagattgactcactcctccagttccagttcaagacacacacttctactaacccac 1020
|||||
Db 972 aacgaattgattgctagattgactcactcctccagttccagttcaagacacacacttctactaacccac 1031
|||||
Qy 1021 actttggaacttaaccacagctacttccattgaaacgctactttgtacgctgaacttctct 1080
|||||
Db 1032 actttggaacttaaccacagctacttccattgaaacgctactttgtacgctgaacttctct 1091
|||||
Qy 1081 caagacaacactatgatatactattttctcgttttgggtttgtacaaacggttaccagaacca 1140
|||||
Db 1092 cagacaacactatggtttctattttcttctcgttttgggtttgtacaaacggttaccagaacca 1151
|||||
Qy 1141 ttgtctactacttctgttgaatctattgaagaaactgacggttactctcgttcttctggact 1200
|||||
Db 1152 ttgtctactacttctgttgaatctattgaagaaactgacggttactctcgttcttctggact 1211
|||||
Qy 1201 gttccattcgtcgtgagacttacgttgaatgatgcaatgtcaagctgaaaggaacca 1260
|||||
Db 1212 gttccattcgtcgtgagacttacgttgaatgatgcaatgtgaaaggaacca 1271
|||||
Qy 1261 ttggttagagttttgttgaacacagaggtgttccattgacaggttgtgtgttgacaag 1320
|||||
Db 1272 ttggttagagttttgttgaacacagaggtgttccattgacaggttgtgtgttgacaag 1331
|||||
Qy 1321 ttggttagagttgaagagacgactcgttgaaggtttgtcttctcgtcgtagatcgtgtgt 1380
|||||
Db 1332 ttggttagagttgaagagacgactcgttgaaggtttgtcttctcgtcgtagatcgtgtgt 1391
|||||
Qy 1381 aactgggctgaatgtttcgtctaa 1404
|||||
Db 1392 aactgggaagaatgtttcgtctaa 1415
|||||

QY 181 tctgtatttccagagcgttccagagcgttagagttactttcgttcaagtttgtct 240
Db 181 TCGGTTCATCTCCCTAGGTGCGCGCGGATGAGATGCTACTTTCGCTCAGGTCCTCC 240
QY 241 agacacggtgctagataccacaaacttcttctgctctaaagcgttactctgtttgattgaa 300
Db 241 GTCATGAGCGCGGTATCCGACCGAGTCTCAAGGCGAAGAAATACTCCGCTCATTTAG 300
QY 301 gctattcaaaagacgctactgtttcaaggttaagtacgcttcttctgaagacttaaac 360
Db 301 GAGATCCAGCAGACGCGACCACTTTGACGGAAATATGCTTCCGTAAGACATACAC 360
QY 361 tacactttggtgctgacgacttacttccattcgttgtaaaccaaatggttaactctggt 420
Db 361 TACAGCTTGGGTGAGATGACCTGACTCCCTTCGGAACAGGAGCTAGTCAACTCCGGC 420
QY 421 attaatctacagaagcgttctgctgctgctgctgctgctgctgctgctgctgctgct 480
Db 421 ATCAAGTCTACAGCGGTAGGATGCTCACAAGGAACATCTCTCATTCATCCGATCC 480
QY 481 tctgtttctacagaggtattgtcttctgctgctgctgctgctgctgctgctgctgct 540
Db 481 TCTGCTCCAGCGCGGTGCTGCTCGGCAAGAAATTCATCCAGGCTTCCAGACACC 540
QY 541 aagttgctacccaggttctacacacacacacacacacacacacacacacacacacac 600
Db 541 AAGCTGAAGGATCTCTGCTGCGCCAGCGCGCAATCTCGCCCAAGATCGACGTGTCATT 600
QY 601 ccagaaggttcggttacaacacactttgaccacggttactgttactgttcttcaagac 660
Db 601 TCCGAGCCAGCTCATCCACACACTCTGACCCAGGACCTGCACTGCTTTCGAAGAC 660
QY 661 tctgaattaggtgacagcgttgaagtaacttcaactgcttcttctgctccagctattaga 720
Db 661 AGCGAATTGCGCGATACCGTCGAAGCAATTCACCGCCACGCTGCTGCCCTCCATTCT 720
QY 721 gctagattgaagcactacacttgcaggtgcttacttgaactgaagacgcttcttactgt 780
Db 721 CAAGCTGGAGAACGACGCTGCTCGGTGCTGCTCTCACAGACACAGAGTACCTACCTC 780
QY 781 atggacatgttccattcagacgttcgctgagaaacttctgacgctactgaattgtctca 840
Db 781 ATGGACATGTGCTCTTGACACCACTCTCCACGACCGCTGACACCAAGCTGCCCC 840
QY 841 tctgtgcttctactcagcagcgaatggtatccaatagcacttctgcaagcgttgggt 900
Db 841 TCTGTGACTGTGTCACCATGCAATGATCACTACCTACCTACCTCAGGCTCTTGA 900
QY 901 aagtaactcgtttacggttgcgttaaccattgggtcagcgtcaaggttgggttgcgt 960
Db 901 AAGTATTACGGCCATGGTGCAGGTAAACCGCTCGGCGCGACCCAGGGGCTCGGCTACG 960
QY 961 aacgaattgattgctagattgactcacttccagttcagttcaagacacacttctactaacac 1020
Db 961 AACGAGCTCATCGCGCTGTGACCCACTGCGCTGTCCACGATGACACCACTTCCAAACAC 1020
QY 1021 acttggactctaacccagcacttctccattgaacgctatttgaactgacttctct 1080
Db 1021 ACTTTGACTCGAGCCGCGCTACCTTTCCGCTCACTTACTCTACTCGCGGACTTTTCG 1080
QY 1081 cagcaacacactatgatctcttcttcttcttcttcttcttcttcttcttcttcttctt 1140
Db 1081 CATGACACGGCATCATCTCCATCTCTTTGCTTTAGGTCTGTACACGGCACTAAGCG 1140
QY 1141 ttgttactacttcttcttgaattctattgaagaactgacggttactctgcttcttggact 1200
Db 1141 CTATCTACACACCGCTGGAGAAATATCACCCAGACAGATGATCTCTGCTCTGAGC 1200
QY 1201 gttcattcgtctgtagagcttactgtaaatgctgaatgctgaatgctgaatgctgaatg 1260
Db 1201 GTTCCGTTTGTCTCGCTTTGTAGCTTCAGATGATGCACTGCTAGGCGGAGCAGGCG 1260
QY 1261 ttggttagagtttggtttaacgacagagattgttctccattgacggttgtgtgttgacaag 1320

Db 1261 CTGGTCCGCTGTCTTGGTTAATGATCGCTGTGCGGCTGCATGGGTGTCGGTTCATGCT 1320
QY 1321 ttgggttagatgaagagacgacgacttctgctgaaggtttgtcttctgctagatcttggt 1380
Db 1321 TTGGGAGATGTACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTCGGGGT 1380
QY 1381 aactgggctgaattgttctgctta 1403
Db 1381 GATTGGCGGAGTGTTCCTTA 1403

RESULT 3

US-08-146-424-19
; Sequence 19, Application US/08146424
; Patent No. 5593963
; GENERAL INFORMATION:
; APPLICANT: VAN OIJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/146,424
; APPLICATION NUMBER: 33,407
; FILING DATE: 02-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KENNEDY, BILL
; REGISTRATION NUMBER: 33,407
; REFERENCE/DOCKET NUMBER: 44615-20011.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1401
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 70
; US-08-146-424-19

Query Match 45.2%; Score 635; DB 1; Length 1404;
Best Local Similarity 65.8%; Pred. No. 5.8e-190;
Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;
QY 1 atggcggtgtctgctgctactgtccattgcaactgttctggttccacatccggtacc 60
Db 1 ATGGCGGTCTCTGCTGTCTACTCTCTTGTATCTCTCTGTGAGTACCTCCGGACTG 60
QY 61 gcttgggtctctggtgtaattctcactcttgtgacactgttgacggtgtgttaccatgt 120

Db 61 GCAGTCCCGGCTCGAGAAATCAATCAGTTGGATACGTCGATCAGGGGTATCAATGC 120
Qy 121 ttccagagaatttctacttgggttacactctccataacttcttcttggcagacgaa 180
Db 121 TTCTCCGAGAGATTCCGATCTTTGGGTCATATACGACCGTTCCTCTCTGCGAAACGAA 180
Qy 181 tctgtatttccagagcttccagagacttagagttacttcttcttccagtttctgtct 240
Db 181 TCAGTATCTCCCTGAGTGGCGCGGATGAGAGTCACTTTGGCTCAGGTCCTCTCC 240
Qy 241 agacaggtgttagatcccaacttcttctgctctaaggtcttacttctgttattgaa 300
Db 241 CGTATGAGCGCGGTATCCGACCGATCCCAAGGGCAAGAAATATGCTTCCTGAAGACATACAAC 360
Qy 301 gctattcaaaagacgtactgtcttcaaggttaagtaagcttcttgaagacttacaac 360
Db 301 GAGATCCAGCAAGCGCACCATCTTGACGGGAAATATGCTTCCTGAAGACATACAAC 360
Qy 361 tacactttgggtgctgacgacttactccatttgggtgaaacccaaatgggttaactctgt 420
Db 361 TACAGCTTGGGTGAGATGACTTCCCTTCGAGAACAGGAGCTTAGTCAACTCCGGC 420
Qy 421 attaatgtctacagagatacaaggttggctagaagaatttcttccattcattagact 480
Db 421 ATCAAGTTCTACACGCGTACGAATCGCTCAAGGAACATFCGTTCCTATCATCCGATCC 480
Qy 481 tctgttctgacagagttattgtcttctgctgaagttcattgaagtttccaaactgtct 540
Db 481 TCTGCTCCAGCGGCTGATCGCTCCGGCAAGAAATTCATCGAGGGCTTCCAGAGCAC 540
Qy 541 aagttgggtgctgacaggttctcaaccacacacagcttctcagtttataaagctgact 600
Db 541 AAGCTGAAGATCTCTGCTGCCAGCGCGCAATCGTGGCCCAAGATCGAGCTGGTCAAT 600
Qy 601 ccagagatccggttcaacacacttggacacaggttactgtactgcttctgaaagac 660
Db 601 TCCGAGGCCAGCTATCAACAACTCTCCAGCCAGCACCTGCACGTGTCTCGAAGAC 660
Qy 661 tctgaattaggtagacaggttgaagtaacttcaactgcttcttcttctcagctattaga 720
Db 661 AGCGAATTTGGCGATACGCTCGAAGCAATTTACCGCCAGCTGCTGCCCTCCATCGT 720
Qy 721 gctagattgaagctgacttccaggttacttctgactgacgagagctgttacttctg 780
Db 721 CAACGCTGGAGACGACTGTCGGGTGTGACTCTCAGACACAGAACTGACCTACCTC 780
Qy 781 atggacatgttccattgacactgctgtagaacttctgacgtactgaattgtctcca 840
Db 781 ATGGACATGTCTCTCGACACCATCTCCACAGCACCGTCCGACACCAAGCTGTCGCC 840
Qy 841 ttctgtcttcttcaactcaagacgaatggatcccaatcagactacttcaagcttgggt 900
Db 841 TTCTGTGACTCTTCAACCATACGAATGGATCAACTACGACTACTCTCCAGCTCTTGA 900
Qy 901 aagtaactcaggttaggtgctgtaacccattgggtcagctcagaggttgggttctgct 960
Db 901 AAGTATTACGGCATGCTGACGATACCGCTCGGCCCGACCCAGGGGTGCGGTACCT 960
Qy 961 aacgaattgattgttagtacttacttccattcctcagttcagacacacacttctactaac 1020
Db 961 AAGGAGCTCATCGCGCTGTGACCCACTCGCGTGTCCAGGATGACACCAAGTTTCAAC 1020
Qy 1021 acttggacttaacccagctacttctccattgaagcttacttcttctgacttctct 1080
Db 1021 ACTTTGGATCGAGCGCGCTACTTCTTCGCTCAACTTACTCTCTACCGGACTTTTCG 1080
Qy 1081 cagcaacacactatgatatacttcttctcgttcttgggtttgttacaacaggttaccag 1140
Db 1081 CATGACAACGGATCATCTCCATCTCTTGTCTTAGGCTCTGACAAACGGCACTAAGCG 1140
Qy 1141 ttgtctactacttctgttgaatctattgaagaactgacggttacttcttcttggact 1200

Db 1141 CTATCTACCACCGCTGAGATATATCACCAGACAGATGATTTCTGCTGCTTGGACG 1200
Qy 1201 gttcattctcgtctagagcttgaattgaatgatgaatgaatgaatgaatgaatgaatga 1260
Db 1201 GTTCGGTTGCTTCGCTTGTACGTCGAGATGATGACAGTGTCCAGCGGAGACGAGCG 1260
Qy 1261 ttggttagatttgggttaacgacagagtttccattgacaggttctgctgttgaacaag 1320
Db 1261 CTGGTCCGCTGCTTGGTTAATGATCGCGTGTTCGCGCTGATGGGTGCTCGGTTGATGCT 1320
Qy 1321 ttgggttagattgaagagagacgacttctgaaggttcttcttcttcttcttcttcttct 1380
Db 1321 TTGGGAGATGATACCGGATAGCTTTGCTGAGGGGTGAGCTTTGCTAGATCTTGGGGGT 1380
Qy 1381 aactgggtgaatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1403
Db 1381 GATTGGCGGAGTGTTCCTTTA 1403

RESULT 4
US-08-693-709-1
; Sequence 1, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OIJEN, ALBERT J.J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1401
; OTHER INFORMATION:
; NAME/KEY: mat.peptide
; LOCATION: 70...1401
; OTHER INFORMATION:

NAME/KEY: Signal Sequence
 LOCATION: 1...72
 OTHER INFORMATION:
 US-08-693-709-1

Query Match 45.28; Score 635; DB 1; Length 1404;
 Best Local Similarity 65.88; Pred. No. 5.8e-130;
 Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;

```

QY 1 atggcggtgttcgtcgtctactgtccattgccacattgttcgttcacacatccggtaacc 60
Db 1 atggcggtgttcgtcgtctactgtccattgccacattgttcgttcacacatccggtaacc 60

QY 61 gcttgggtgttcgtcgtctactgtccattgccacattgttcgttcacacatccggtaacc 120
Db 61 gcttgggtgttcgtcgtctactgtccattgccacattgttcgttcacacatccggtaacc 120

QY 121 tcccgagaaattctcacttgtgggtgacctactctccatctctccatctctcttggcagacgaa 180
Db 121 tcccgagaaattctcacttgtgggtgacctactctccatctctccatctctcttggcagacgaa 180

QY 181 tctgctatttctccagacgttccagacgtgttagagttactctctccatctctcttggcagacgaa 240
Db 181 tctgctatttctccagacgttccagacgtgttagagttactctctccatctctcttggcagacgaa 240

QY 241 agacacgggtgttagatccacaaattctctgctctgaaggtctactctctcttggcagacgaa 300
Db 241 agacacgggtgttagatccacaaattctctgctctgaaggtctactctctcttggcagacgaa 300

QY 301 gctattcaaaagacgtactgttctcaagggttaagtaagtaagtaagtaagtaagtaagtaagtaag 360
Db 301 gctattcaaaagacgtactgttctcaagggttaagtaagtaagtaagtaagtaagtaagtaagtaag 360

QY 361 tacacttgggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 420
Db 361 tacacttgggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 420

QY 421 attaatgtctacagaagatacaagggttgggtgagaaagatttccattcattagagct 480
Db 421 attaatgtctacagaagatacaagggttgggtgagaaagatttccattcattagagct 480

QY 481 tctgttctcagacagattgttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 540
Db 481 tctgttctcagacagattgttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 540

QY 541 aagttggtcaccacaggttctcaccacacacacacacacacacacacacacacacacacacacacacac 600
Db 541 aagttggtcaccacaggttctcaccacacacacacacacacacacacacacacacacacacacacacac 600

QY 601 ccagaagattccggtttacacacacacacacacacacacacacacacacacacacacacacacacacacac 660
Db 601 ccagaagattccggtttacacacacacacacacacacacacacacacacacacacacacacacacacacac 660

QY 661 tctgaattagtgacagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttga 720
Db 661 tctgaattagtgacagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttga 720

QY 721 gctagattggaagtgactgtccaggtgttactgttgaagcgttgaagcgttgaagcgttgaagcgttga 780
Db 721 gctagattggaagtgactgtccaggtgttactgttgaagcgttgaagcgttgaagcgttgaagcgttga 780

QY 781 atggacattgttgcattgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcac 840
Db 781 atggacattgttgcattgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcac 840

QY 841 tctgtgttcttgcattgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcac 900
Db 841 tctgtgttcttgcattgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcacagcgttgcac 900

QY 901 aagtaactcaggttgcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 901 aagtaactcaggttgcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960

```

RESULT 5

US-08-419-448-33
 ; Sequence 33, Application US/08419448
 ; Patent No. 5863533
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert F.M. Van Gorcom
 ; APPLICANT: Willem Van Hartingsveldt
 ; APPLICANT: Petrus A. Van Paridon
 ; APPLICANT: Annemarie E. Veenstra
 ; APPLICANT: Rudolf G.M. Luttin
 ; APPLICANT: Gerardus Selten
 ; TITLE OF INVENTION: Cloning and Expression of Microbial
 ; TITLE OF INVENTION: Phytase
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morrison & Foerster
 ; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1888
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/419,448
 ; FILING DATE: 10-APR-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20026.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-887-1500
 ; INFORMATION FOR SEQ ID NO: 33:
 ; SEQUENCE CHARACTERISTICS:

Qy 613 ggttaacacacactttggaccacggttaactgtactgtcttcaaacactctgaattaggt 672
Db 1118 TCATCCAAACACTCTCGACCCAGGACCTGACACTGCTTTGAAGACACGCAATTTGGCC 1177
Qy 673 gacgacgtgaagtaacttaactcactgttttgcctccagctattagagtagatgaa 732
Db 1178 GATACCGTCGAAGCCANTTTCACGCCACAGTTCGCCCCCTCCATTGCTCAACGTCGGAG 1237
Qy 733 gctgacttgcaggtgttaactttgactgaagaaagacgttgtttacttgatggacatgtgt 792
Db 1238 AAGACCTGTCTGCGTGACTCTCACAGACACAGAAAGTACCTACCTCAATGGACATGTC 1297
Qy 793 ccattcgacactgtcgttagaacttctgaagcactactaaatgtctcattctgtgtttg 852
Db 1298 TCCCTTCGACACCATCTCCACAGCACCGTGCACCAAGCTGTCCCTCTCTGTGACCTG 1357
Qy 853 ttactacacagcaatgagatcccaatacagactacttgcacagcttgggtaagtactacggt 912
Db 1358 TTCACCCATGAGAAATGATGATCCACTAGACTACCTCCAGTCCCTGAAAAAATACTACGGC 1417
Qy 913 tacgggtgctgaacacactttgggtccagctcaaggtgttggtttcgctaacgaattgatt 972
Db 1418 CATGGCGAGGTAACCCGCTCGGCCCGACCGAGGCTCGGCTACGCTAACGAGCTCATC 1477
Qy 973 gctagattgactcactccagttcaagacacacacttctactaacacacacttttgactct 1032
Db 1478 GCCGCTCACCACTCCGCTGTCACGATGACACAGCTCCACACACACTTGGACTCG 1537
Qy 1033 aaccagactacttccacttgaacgctactttgtacgctgacttctctcaacacacact 1092
Db 1538 AACCCAGCTACTCCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAAGGGC 1597
Qy 1093 atgatatatttttctgtgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgt 1152
Db 1598 ATCATCTCTATCTCTGCTTTGGGTTCTGTACACAGGCAATAAGCCGCTGTCTACACAG 1657
Qy 1153 tctgtgaactatgaagaactgacggttactctgtctgttggactgttccatctogct 1212
Db 1658 ACCGTGGAGATATACCCACAGACAGATGGTTCGCTGCTGTGGACGGTTCGCTTTGCT 1717
Qy 1213 gctagagcttaacttgaatgatgaatgtcaagctgacggttgcgtgttgacaaagtgggtagatgt 1272
Db 1718 TCGCGTCTGTAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
Qy 1273 ttggttaacgacagaggtttgttccatgtcacggttgcgtgttgacaaagtgggtagatgt 1332
Db 1778 TTGGTTAATGATCGCGTGTCTCCGCTGATGGGTGATGGGTGATGGGTGATGGGTGATGG 1837
Qy 1333 aagagacgacttctgttgaaggtttgttcttcgctagatctgtgttaactgggctgaa 1392
Db 1838 ACCCGGATAGCTTTGTAGGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTTGGGCGGAG 1897
Qy 1393 tottgcgtta 1403
Db 1898 TGTTCTGCTTA 1908

RESULT 9
US-08-609-426A-7
Sequence 7, Application US/08609426A
Patent No. 5830733
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloheimo, Marja T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Rambosek, John A.
APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
APPLICANT: Houston, Christine S.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes

TITLE OF INVENTION: in Trichoderma
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,426A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,724
FILING DATE: 31-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Grant E.
REGISTRATION NUMBER: P-41,264
REFERENCE/DOCKET NUMBER: 1050.0080001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2363 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: Join(404..447, 550..1906)
US-08-609-426A-7

Query Match 44.8%; Score 629.4; DB 2; Length 2363;
Best Local Similarity 65.8%; Pred. No. 4.4e-188;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
Qy 13 gtcgtgctactgtccattgccacactgttgcgttccacacatcccggtaccgccttgggtcct 72
Db 518 GTGGGACTACTGATCGTGACAAATCTGTGCAGAGTCACCTCCGGACTGGCAGTCCCGGCC 577
Qy 73 cgtggttaattctcactcttctgtgacactgttgacggttggttaccaaatgtttccagaaatt 132
Db 578 TCGAGAAATCAATCCACTTTCGATACGGTGCATCAAGGGGTATCAATGCTTCTCCGAGACT 637
Qy 133 tctcactgtgggttacctactctccatacttcttcttgacagacgaactctgctattct 192
Db 638 TCGCATCTTTGGGGTCAATACGCGCCCTTCTCTCTCTGCAACACGAATCGGCCATCTCC 697
Qy 193 ccagacgttccacacgactgttagatttacttctggttctaaagttttgtctagacacggtgct 252
Db 698 CCTGATGTGCCCGCGGTTGCAGAGTCACCTTTCGCTCAGGTCTCTCCGCTCATGGAGCG 757
Qy 253 agatacccaacttcttctggttgaaggttactgcttctgattgaagctattcaaaag 312
Db 758 CGGTATCCGACCGAGTCCAAAGGGCAAGAAATACTTCCGCTCTCTATTGAGGAGATCCAGCAG 817

QY 313 aacgactgcttttaagggtaagtaagcttttttaagactttacaactacactttgggt 372
DB 818 AACGTGACCACTTTGATGAAATATGCTTCTGAAGACATACACTAGCTTGGGT 877
QY 373 gctgaacgactgactcattcgggtgaacacacaaatggttaactctggtattagttctac 432
DB 878 GCAGATGACCTGACTCCCTCGGAGACGAGAGTAGTCAACTCCGGCATCAAGTTCTAC 937
QY 433 agaagatacaaggctttggctagaagatgttccattcattagagcttctggttcgac 492
DB 938 CAGCGATAGCAATCGCTACAGGAACATATTCCTGTCATCCGATCCTCTGGCTCCAGC 997
QY 493 agattattgcttctgctgaagaaatggttaagaggtttccaaatgctgaagttggtgac 552
DB 998 CGCTGATGCTGCTCGGCGAGAAATTCATTGAGGCTTCAGAGACACCAAGCTGAAGGAT 1057
QY 553 ccaggttctcaacacacacacaggtttccaggtatttaacgtgatcattccagaagatcc 612
DB 1058 CTGCTGCCCGCGGCCCAATCTCGCCCAAGATCGATCGTGTCAATTCGAGGCGCAGC 1117
QY 613 ggttacacacacactttggaccacggttactgtactgtcttcgaagactctgaattaggt 672
DB 1118 TCATCCAAACAACACTCGACCCAGGACCTGCACTCTTTGAAGACACGGAATTGGCC 1177
QY 673 gacagactgaagacttaacttactgtttgttgcctccagctattagagcttagattggaa 732
DB 1178 GATACCTCGAAGCAATTCACGCCACCTTCGCCCCCTCCATTCGTCACGCTCTGGAG 1237
QY 733 gctgaactgcaaggtgttaactgtactgacgaagaagcgtttgttaactgtatgacatgtgt 792
DB 1238 AACGACCTGTGCGGTGACTCTCACAGACACAGAAAGTACCTTACCTCATGGACATGTC 1297
QY 793 ccattgacacactgctgtagaacttctgacgtactgaattgttccattctgtgtgttg 852
DB 1298 TCCTTCGACCACTCTCCACGACGCTCGACCAAGGTGTCGCCCTCTGTGTGACCTG 1357
QY 853 ttoactcacgaacgaatggtatccactacgactacttgcgaagcttgggttaagtactacgt 912
DB 1358 TTACCCCATGCAATGGATCCACTAGGACTACTTCCAGTCCCTGAAATAACTACGGC 1417
QY 913 tacggtgctgaacacacattggtcagctcgaaggtgtgttgcctgacgaattgatt 972
DB 1418 CATGGCGCAGGTAAACCGCTCGCGCGGCCGCCAGGGCGTGGCTTACGCTAACGAGCTCATC 1477
QY 973 gctagattgactcactctcagttcagacacacacttctactaacacacacttggactct 1032
DB 1478 GCCGCTTACCCACTCGCTGTCCACGATGACCCAGCTCCACACACACACTTGGACTCG 1537
QY 1033 aacccagctacttccattgaacgctactttgtacgctgacttctctcacgacacact 1092
DB 1538 AACCCAGCTACCTTCCCGCTCAACTCTACTCTAGCGGACTTTTCCACCATTAACGGC 1597
QY 1093 atgatactattttctgcttgggtttgtaacagcgttaccagcacttggcttactact 1152
DB 1598 ATCATCTACTCTCTTGTGCTGTGTACACAGGCTTCCACCATTAACGGC 1657
QY 1153 tctgtgaactattgaacaaactgaacgttactctgttcttggactgttccattcgtct 1212
DB 1658 ACCGTGGAGAAATACCCAGACAGATGGGTCTCTGCTGTGCTTGGACGTTCCGTTTGT 1717
QY 1213 gctagagcttgaattgaatgataatgataatgataatgataatgataatgataatgatt 1272
DB 1718 TCGGCTCTAGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
QY 1273 ttggttaacgacagagttgttccattgacggttgcctgttgcctgttgcctgttgcctgt 1332
DB 1778 TTGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1837
QY 1333 aagaagacgacttctgtgaaggtttgttcttgcctgtagatctgtgttaactgggtgaa 1392
DB 1838 ACCCGGATAGCTTGTGAGGGGGTTGAGCTTTGCTAGATCTGAGGGGTGATTTGGCGGAG 1897
QY 1393 tgtttcgcctta 1403

Db 1898 TGTTCTGCTTA 1908

RESULT 10

US-08-374-652C-1
Sequence 1, Application US/08374652C
Patent No. 5834286
GENERAL INFORMATION:
APPLICANT: NEVALAINEN, HELENA K.M.
APPLICANT: PALOHEIMO, MARJA T.
APPLICANT: FAGERSTROM, RICHARD B.
APPLICANT: MLETTINEN-OINONEN, ARJA S.
APPLICANT: TURUNEN, MARJA K.
APPLICANT: RAMBOSEK, JOHN A.
APPLICANT: PIDDINGTON, CHRISTOPHER S.
APPLICANT: HOUSTON, CHRISTINE S.
APPLICANT: CANTRELL, MICHAEL A.
TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
VECTORS AND METHODS FOR EXPRESSING PHYTATE DEGRADING
TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07058
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2379 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(404..447, 550..1906)
US-08-374-652C-1

Query Match 44.8%; Score 629.4; DB 2; Length 2379;
Best Local Similarity 65.8%; Pred. No. 4.5e-188;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
QY 13 gtagtgactgcttgcattgccaactgttgccttgccttgccttgccttgccttgccttgcct 72
DB 534 GTGGGACTACTGATCGCTGACAACTGTGCGAGAGTCACTTCCGGACTGGCGAGTCCCGCC 593

Db 209 cgggtcgatcagggtatcaatgcttctcggagagacttcgcatctttggggccaatacgcgc 268
QY 158 catacttctcttggcagacgaatctgctatttctcagagcttccagagactgttagag 217
Db 269 cgttcttctctcgcgaacaaatcggccatctccctctgagtctcctgcooggatgcgatg 328
QY 218 taacttctcgttcaagtttctgtagacacggtgctagatacccaacttcttctcgtctta 277
Db 329 tcaacttgcgccaggttctctcccgccatggagcacggtatccgacgactccaagggca 388
QY 278 aggtctactctgcttctgattgaactattcaaaagacgctactgcttccaaggttaagt 337
Db 389 agaaataactcgcctctcatcgagagatccagcagaacgagcaactctcggagggaaat 448
QY 338 acgcttcttgaagacttaacactacacttgggtgctgacgacttgactccattcgggtg 397
Db 449 atgcttctcgaagatacaactacagctcgggctggatgactgactccctctcggag 508
QY 398 aaaacaaatgggttaactctggtatttaagtctacagaagatacaagcttgggttagaa 457
Db 509 agcaggagctggtcaactcgggtcgaagttctaccagcgatagcaatcgctcacaaaga 568
QY 458 agattgtccattcattagagcttctgttctgacagagttattgcttctcgtcgaagaagt 517
Db 569 acattgtccgcttccatccgactcccgaggtccagccgctgattgctctcgtcgaataat 628
QY 518 tcattgaaggtttccaatctgctaagtgtggtgacccaggttctcaacacacacacactt 577
Db 629 tcactcggaggttccagagcactaagtgatcctcgtgcccagcccgcccaactcgt 688
QY 578 ctccagttataacgtgactcattccagaaggtatccggttacaacacacacttggacacg 637
Db 689 cgcccaagatcgctggtcatttcagagcgccagcactccacacacactctcgatccgg 748
QY 638 gtacttactgcttctcgaagactcgaattgagtgacagcttggaagctaaactcactg 697
Db 749 gcaactgcacggttttcgaagatagcgaattggccgactgacatcgagccaatttcacg 808
QY 698 cttgttcgtccagctatttagagctagattggaagctgacttgcaggtgttactttga 757
Db 809 ccaagttctccctccactcgtcgaactggtgagacgacttgtctgctgctctca 868
QY 758 ctgacgaagacgtgttactgtaggacatgtgtccattcgacactgtcgtagaactt 817
Db 869 cggacacagaagtgactccctcatgacatgtgctccttcgacacactctccacagca 928
QY 818 ctgacgactgactgaattgtccattctgcttctgttctcactcagcagaatgactccaat 877
Db 929 cgtcgacacacagctgccccctctgtgactgttccaccatgaagaatggatcaact 988
QY 878 acgactacttgaaagcttgggttaagtagtactacggttaccggtgctgtaacccattgggtc 937
Db 989 acgactactccagtcctcctgaacaaatactacggtccatggcgaggttaacccgctggcc 1048
QY 938 cagctcaaggtgtgttctcgtacgaattgattgctagattgactcactcactccagttc 997
Db 1049 cgaccagggctcggctacgctacagagctcatcgccgtctcaccactcgcgtctgcc 1108
QY 998 agagacacactcttactaaccacacttggacttaaccacagacttctccattggaagc 1057
Db 1109 acgatgacaccagctccaacacacattggactccaacccggtacttcccgctcaact 1168
QY 1058 ctactttgactgacttctctcagcaacacactatgatatactatttcttctcgtttgg 1117
Db 1169 ccactctctatcgacttctcgtatgataacggtatcactctcactctcttcttcttgg 1228
QY 1118 gttgtacacgggttaccagccattgtctactactctgttgaaactatttgaaagaaactg 1177
Db 1229 gctgtacacggcacaagccgctgtcttccagaccgctggagaaatatacccccagaccg 1288
QY 1178 acggttactctgtcttcttggactgttccattcgtcgttagagcttaacgttgaatgatgc 1237
Db 1289 atgggttctcactcgtcgtgacggttctcttctcgtcgtcgtatgactgacgtcagatgatgc 1348

QY 1238 aatgtcaagctgaaaaggaaaccattggttagagtttgggttaacacagagttgttccat 1297
Db 1349 aatgcccagtcgcagcagagagccttggtcggtcttggtaataatgctgttgttccgc 1408
QY 1298 tgcacggttctgctgttgacaagattgggttagatgttaagagagacgactcgttgaaggtt 1357
Db 1409 tgcattgctcgcggttggatgcttgggaagatgtacgcgggatgcttctgtaaggggt 1468
QY 1358 tgcatttctgtagatcgttgggttaactgggtgagatgttctcgttca 1403
Db 1469 tgaagtttccagatcgtcgggtgattggggggaggtcttctcgttca 1514

RESULT 12

US-09-155-855-4
; Sequence 4, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidekazu
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ. ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1332)
US-09-155-855-4

Query Match 43.5%; Score 610.2; DB 3; Length 1332;
Best Local Similarity 66.2%; Pred. No. 3.6e-182;
Matches 879; Conservative 0; Mismatches 448; Indels 0; Gaps 0;

QY 77 gtaattctcactctgtgacactgttgacggtgttaccatgtttcccaatgtttcccaagaattctc 136
Db 5 gaatacattccacttgcgtacgacggtcaggggtatcattcgtctcgtcggagacttcgc 64
QY 137 acttgtgggttacctctctccactctcttcttggcagacgaatctgctatttctccag 196
Db 65 atcttggggcacaatacgcgcgcttcttctctggaacaaatcgccactctccctg 124
QY 197 acgttcagacgacgtgtagagttacttctgcttaagtttcttagacacggtgtagat 256
Db 125 atgtctctcgggagtcactgtcacttctccaggttctctcccgccatgggagcaggt 184
QY 257 acccaactctctcgtcttaaggttactctgcttcttggattgaagctattcaaaagacg 316
Db 185 atccgaccgactccaaaggcagaataactcctctctcctcgtcggagagccagagacg 244
QY 317 ctactgcttcaaggggtaagtagcgttcttgaagacttacaactacacttttgggtgctg 376
Db 245 cgacaacttcgaggggaaatagcttctctgaacatacaactacagcctggcgcg 304
QY 377 acgattgactcattcgttgaaacaaatggttaactctggtatttaagttctacagaa 436
Db 305 atgactgactcctcctcggagagcaggtggtcactccgctcaggttctaccagc 364
QY 437 gatacaggtcttggctagaaagattgttccattcattagagcttctggttctgacagag 496


```
QY 607 ggaatccggtttacaacaacacactttggaccacggtactgtactgtgttgcgaagactctgaa 666
Db 1087 GAACTGGCAGTAACAACTACTCTGATGGCTGACGTGCCCGCGCGGAGGAGGACCGG 1146
QY 667 ttaggtgacgagcttgaagacttaacttctactgtctgttgcctccagctattagagctaga 726
Db 1147 GACCCAAACCCAGCCGACAGTTCCTGCAAGTTTTCGGCCCGCGTGTCTGNAAGATC 1206
QY 727 ttggaagctgacttgcaggtgttactttagtgcagcaggaagcgttgttactttagtgac 786
Db 1207 ACTAAA---CACATGCGGGGTGTAACCTACCTTTGGAGGATGTCCCGTTGTTCATGAT 1263
QY 787 atgtgtccattgacactgt-----cgttagaactctctgacgtactgaattgtctcca 840
Db 1264 CTTTGTCCGTTGACACGGTGGCTCCACACCGAGTCTTTTCCACGGCAGCTCTCTCCG 1323
QY 841 tctgtgttcttcaactcaagcagaaatggatccaatacagactacttgcgaagcttgggt 900
Db 1324 TTTTGTCACTTTTTCACGGCCGACGATTGGATGGCTACGATTACTACTACACCTCGAC 1383
QY 901 aagttactcgtttagcgtgctggttaacccattgggtccagctcgaaggtgttgggttgcgt 960
Db 1384 AAATACTACAGCCAGCGCGGCGGACGCAATTTGGCCGCTCCCGCGCGTGGGTTCGTC 1443
QY 961 aacgaattgattgtagattgac-----tcaactctccagttcaagcaccacacttctactaac 1017
Db 1444 AACGAGCTGATTGCGCGTATGACGGGAATCTTCCCGTCAAGGACCCACACAGTCAAC 1503
QY 1018 cacacttggactctaaaccagctacttcccaattggaacgacttcttgcagctgaattc 1077
Db 1504 CACACTCTCGATGACAAACCGGAACTTTCCGTTGGACGCTGTCTCTACGACAGACTTT 1563
QY 1078 tctcagcacaactatgatatctatttcttgcgttgggttgggttgcagcaggttaccag 1137
Db 1564 TCGCAGCAGACACACCACTGACGGGCACTTTTCCGCAATGGGCCCTGTACACGGCACAAAG 1623
QY 1138 caattgtctacttctgttgaattct-----attgaagaactgacggttactctgct 1191
Db 1624 CGCTGTGCGAGTCCAAAGATTGAGCTCCGACGGGTGCAGCAGCGGATGATATGCGGCA 1683
QY 1192 tcttgcagcttccattcgtctagagcttgaagcttgaagcttgaagcttgaagctgaa 1251
Db 1684 TCGTGGAGGCTGCGCTGTCAGCAGCGGCTATGTGAGATTGTGCGATGTGACAGCGAA 1743
QY 1252 aag-----gaaccattggttagagtttgggttaac 1281
Db 1744 ACGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1803
QY 1282 gacagagttgttcattgcaaggttgcgttgcagagttggttagatgtagagagagac 1341
Db 1804 GATCGGTTGTGCGCTGCTATGTTGTGCGGTTGATCGATGGGGAGGTTGTCGAGGGAT 1863
QY 1342 gacttcgttgaaggttcttcttgcagctgagctggttgaactggcgtgaatttt 1397
Db 1864 GASTGGATTAGGAGACTACGTTTGTCTCGACAGGGTGGGCATTGGGATCCTGCTT 1919
```

RESULT 14

```
us-09-163-642-1
; Sequence 1, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 62216440 No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
```

```
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-163-642-1
```

```
Query Match 23.7%; Score 333.2; DB 4; Length 2200;
Best Local Similarity 56.6%; Pred. No. 6.6e-95;
Matches 745; Conservative 0; Mismatches 523; Indels 48; Gaps 5;

QY 127 gaaattctcaacttgggttacctactctccactctctcttcttgcgacagaaactctgt 186
Db 607 GACATTGCCCGCCGACTGGGCGGAGTACTCGCCCTTCTTCGCTGCCGCGAGTCTCTGAA 666
QY 187 atttctccagaagttccagacgactgtagttaaacttctcggttcaagtttctctagacac 246
Db 667 ATCTCGCTGCGGTGCGCAAGGGCTGCTCGTGTGAGTTGTGTCAGGTCTGTGTCGGGAC 726
QY 247 ggtgctagataccacacttctctcgctgaaggcttactctgttcttgcaggttctgt 306
Db 727 GGAGCTCGGTATCTCTACTCTCACAAGAGTGAAGTCTACGCGAGTGTGCTTCAAAGGATC 786
QY 307 caaagaacgctactgttctcaagggttaagtagcgttcttctgaagacttaactaacact 366
Db 787 CAGGACACTGCCACCGAGTTCAGGGCGGATTTTGCTTCTCCGAGACTATGCCCTATCAT 846
QY 367 ttgggtgctgacgacttgcactccattcgtgaaacacaaatggttaactctgttattaa 426
Db 847 CTCGGTCCGCAATAATTGACGCGCTTTGGCAGAGGAGAGATGATGGAATCGGGCCGCG 906
QY 427 ttctcagaagatacaagcgttggctgaaagattgttccattcattagagcttctgtt 486
Db 907 TTCTACCCCGGTATCGGTGAGCAGGCGCGGAGAGATTGTGCGCATTTGTGCGTGGCGAG 966
QY 487 tctgacagagttattgttctctgtgtaaaagtcttcaaggtttcccaactctgtaagt 546
Db 967 TCCGCGCGAGTTCATTGCGTGGCAGAGTTCTTCAACCGCGGATTCAGGATGCCAAGAC 1026
QY 547 gctgaccaggttctcaaccacacacacacacacacacacacacacacacacacacacac 606
Db 1027 CGGGATCCCGAGGTGCAACAAAGAGCAGCAGGAGCAGGCTGTGATCAAGTGATCATTTCC 1086
QY 607 ggtccggtttacaacaacacacttggaccacggttactgtactgttgcgaagactctgaa 666
Db 1087 GAACTGGCAGTAACAACTACTCTGATGGCTGCGGTGCGGCGCGGCGGAGGAGGACCG 1146
QY 667 ttaggtgacgagcttgaagcttaacttctactgttcttgcgtctccagctattagagctaga 726
```

Db 1147 GACCAACCCAGCCGAGAGTCTCTGCAAGTTTTTCGGCCGCGTGTCTTGAAGAAAGATC 1206
Qy 727 ttggaagctgactgcaagtggttacttactgactgcaagagcgttgttacttactgagac 786
Db 1207 ACTAAA--CACATGCCGGGTGTGAACCTCACCTTGGAGGATGTCCCGTGTGTCATGGAT 1263
Qy 787 atgtctcactgacactgt-----cgctagaacttctgacgtactgaattgttccca 840
Db 1264 CTTTGTCTGTTGACACGGTGGGTCTCGACCCAGTCTTTTCCACGGCAGCTCTCTCCG 1323
Qy 841 ttctgtgttttactcaacgagcaatggatccaatacagactacttgcgaagcttgggt 900
Db 1324 TTTTGTCACTTGTTCACGCCGACGAGTGGATGGCTACGATTACTACTACACCCCTGAC 1383
Qy 901 agtactcaggtgactgctggtgtaaaccaattgggtccagctcaaggtgtgtgttgcgt 960
Db 1384 AAATACTACAGCCAGCGGGGGGACGCAATTTGGCCCGTCCCGCGGCGTGGGTCTGCTC 1443
Qy 961 aacgaattgactgattgac---tcactctcagttcaagaccacacttctactaac 1017
Db 1444 AACGAGCTGATTGCGCGTATGACGGGAAATCTCCGTCAGAGGCCACACACAGTCAAC 1503
Qy 1018 cacacttggacttaacccagctacttccattgaaacgtacttgttacgctgacttc 1077
Db 1504 CACACTCTCGATGACACCCCGGAACTTTCCCGTTGGAGCGTCTCTACGCGACTTT 1563
Qy 1078 tctcagcaaacactatgatatacttcttcttcttcttcttcttcttcttcttcttctt 1137
Db 1564 TCGCAGCAGAACACCATGACGGGCACTTTTCCGCAATGGGCGCTGTACACGGGCACAAAG 1623
Qy 1138 ccattgtctactacttcttcttgaatct-----attgaagaaactgacggttactctgct 1191
Db 1624 CGCTGTGACGTCTCAAGATTCAGCTCCGCGGCGTGGAGCAGCGGATGATACCGGCA 1683
Qy 1192 tcttggactgttccattcgtctgtagagcttacttacttgaatgcaatgtcaagctgaa 1251
Db 1684 TCGTGGAGGTGCGCTGTCACGAGGGCGTATGTGGAGTTGCTCGGATGTGAGACGGAA 1743
Qy 1252 aag-----gaaccattgtttagattttagttaac 1281
Db 1744 ACAGAGCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1803
Qy 1282 gacagagttgttccattgcaagcttctgctgttgacaagttgggttagatgtaagagagac 1341
Db 1804 GATCGGGTGTGCGCTGATGTTGTCGGGTGATCGATGGGGGAGGTGTCGGAGGGAT 1863
Qy 1342 gacttctgtaaggttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1397
Db 1864 GAGTGGATTAAAGGAGCTACGTTTGTCTGACAGAGGTGGGCAATTTGGGATCGCTGCTT 1919

RESULT 15
US-09-221-654-1
; Sequence 1, Application US/09221654
; Patent No. 6054306
; GENERAL INFORMATION:
; APPLICANT: Lassen, Soren
; APPLICANT: Bech, Lisbeth
; APPLICANT: Fuglsang, Claus
; APPLICANT: Ohmann, Anders
; APPLICANT: Breinholt, Jens
; APPLICANT: Ostergaard, Peter
; TITLE OF INVENTION: Peniophora Phytase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6054306 No. 6054306disk of No. 6054306th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,654
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/989,358
FILING DATE: 12-DEC-1997
APPLICATION NUMBER: 0529/97
FILING DATE: 07-MAY-1997
APPLICATION NUMBER: 60/046,081
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 5101,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-221-654-1

Query Match 8.7%; Score 122.8; DB 3; Length 1320;
Best Local Similarity 55.7%; Pred. No. 7.8e-29;
Matches 263; Conservative 0; Mismatches 197; Indels 12; Gaps 1;
Qy 663 tgaattaggtgacgacgttgaagtaacttaacttaacttcttcttcttcttcttcttcttctt 722
Db 594 TGAAGTGGATGGTGAAGCAATCCAAACGTTGGGGGTCTTTGGCCGCAACATCACCCG 653
Qy 723 tagattggaagctgacttgcaggtgttactttagctgacgacgctgttacttctgac 782
Db 654 GCGATTGAACGCTGCTGCGCCGAGTGCCAACTCTCAGACAGCAGCGCTCACTCTCAT 713
Qy 783 ggcacatgttccattgacacactgtcgtctagaacttctgacgtactgaattgttctccatt 842
Db 714 GGATATGTGCGCGTTCGACACTCTCAG-----CTCCGGGAACGCCAGCCCTT 761
Qy 843 ctgtgtcttcttactcaacgacgaatggatcccaatacagactacttgcgaagcttgggtaa 902
Db 762 CTGTGACCTATTACCGCGGAGGAGTATGTGTCTGCTAGAGTACTACTATGACCTCGACAA 821
Qy 903 gtactacggttacggtgtctgttaacccattgggtccagctcaaggtgttgggttctgctaa 962
Db 822 GTACTATGGACAGGCGCCCGGGAACGCTCTCGGTCTCTGTCAGGGGCTCGGATACGTCAA 881
Qy 963 cgaattgattgtagattgactcacttctccagttccagttcaagaccacacttctactaacacac 1022
Db 882 TGAGCTGCTTGCACGCTTGACCGGCCAAAGCGTTTCGAGAGCAGAGCAGACGAAACCGCAC 941
Qy 1023 ttgtgacttaacccagactacttctccatgaacgtacttcttctgactgacttctctca 1082
Db 942 GCTCAGACAGGACCCCTGCAACATTTCCCGCTGAAACCGTACCCGTCTTACGCCGACTTCGCA 1001
Qy 1083 cgacacactatgatatacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1134
Db 1002 TGAATAACACATGTTGCGCCATCTTTGCGGCGCTCGGGCTCTTCAACGCCACC 1053

Search completed: October 26, 2001, 15:20:18
Job time: 574 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1868	75.9	467	3	Q9UWZ7	Q9UWZ7	aspergillus
2	1867	75.8	465	3	Q00092	Q00092	aspergillus
3	1860	75.5	467	3	Q93838	Q93838	aspergillus
4	1839	75.1	467	3	Q9HEQ0	Q9HEQ0	aspergillus
5	1839	74.7	466	3	Q00100	Q00100	aspergillus
6	1761	71.5	466	3	Q00085	Q00085	aspergillus
7	1719	69.8	466	3	Q00096	Q00096	aspergillus
8	1313.5	53.4	487	3	Q00107	Q00107	thelaviaceae
9	344.5	14.0	469	3	Q93846	Q93846	thelaviaceae
10	328.5	13.3	442	3	Q74677	Q74677	thelaviaceae
11	237.5	12.1	463	3	Q60172	Q60172	thelaviaceae
12	203	8.2	460	5	Q9VW72	Q9VW72	thelaviaceae
13	192.5	7.8	467	5	Q96421	Q96421	thelaviaceae
14	189	7.7	451	11	Q35217	Q35217	thelaviaceae
15	186	7.6	481	11	Q922L6	Q922L6	thelaviaceae
16	175.5	7.1	453	5	Q96420	Q96420	thelaviaceae
17	175.5	7.1	487	4	Q9UNW1	Q9UNW1	thelaviaceae
18	174.5	7.1	453	5	Q9W438	Q9W438	thelaviaceae
19	174.5	7.1	487	4	Q95172	Q95172	thelaviaceae

Db 121 YSLGADLTPEGEELVNSGKIFQYVESUTRNPIIRSGSGSRVJASGKKFIEGQST 180
Qy 181 KLADPGSQHPASPVINIIPEGSGYNNITLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLKDPRAQGOSSPKIDVIVSEASSNNITLDPGICTVFEDSELADAVEANFTATFVPTIR 240
Qy 241 ARLEADLPGLVTLDEDDVYVLMDCPFDTVARTSDATELSFPCALFTHDEWIQDYLOSIG 300
Db 241 QRLNDLSGSLTDEVTYVLMDCSEDTISTVTDITSLSPFCDLTHDEWIQDYLOSIG 300
Qy 301 KYGAGNPLGAGGCGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGAGNPLGAGGCGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Qy 361 HDNTMISIFFALGNGTKPLSTTSVESIEETDGYASWSVTPFAARAYVEMMOCAKEP 420
Db 361 HDNGIISILFALGNGTKPLSTTTQVITOTDGFSSAWTVPFAARAYVEMMOCAKEP 420
Qy 421 LVRVLNDRVPLHGCADVCKLGRCKRDFVEGLSFARSGGNWAECEFA 467
Db 421 LVRVLNDRVPLHGCADVCKLGRCKRDFVEGLSFARSGGNWAECEFA 467

RESULT 2
ID 000092 PRELIMINARY; PRT; 465 AA.
AC 000092;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN=ATCC 34625;
RX MEDLINE=97288063; PubMed=9143104;
RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
phatase from the fungus Aspergillus fumigatus";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC EMBL: U59804; AAB96872.1; -.
DR HSP; P34752; 1IHP.
DR InterPro: IPR000560; -.
DR Pfam: PF00328; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 465
FT ACT_SITE 81 81
FT ACT_SITE 82 82
FT ACT_SITE 359 359
FT ACT_SITE 30 39
FT DISULFID 70 412
FT BY SIMILARITY.

FT DISULFID 213 463 BY SIMILARITY.
FT DISULFID 262 280 BY SIMILARITY.
FT DISULFID 434 442 BY SIMILARITY.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 50836 MW; 86FCID9038C9B2C9 CRC64;
Query Match 75.8%; Score 1867; DB 3; Length 465;
Best Local Similarity 76.0%; Pred No. 1.6e-141;
Matches 355; Conservative 36; Mismatches 74; Indels 2; Gaps 2;
Qy 1 MGVEVVLLSIATLFGSTGTLGRNSHSCDTPDGGYOCFPEISHLGWTSPYSLADE 60
Db 1 MVTLTLLSAAYLLSGRVSAAFPSSAG-SKSCDTVDLGYOCSPATSHLWGQYSPFFSLEDE 59
Qy 61 SATSPDVPDCRVTFQVLSRHGARYPTSSAKYASALIEAIQKNATAFKGYAFKLYN 120
Db 60 LSVSKLPKDCRITLVQVLSRHGARYPTSSKSKYKLVTAIOANATDFKGFALKLYN 119
Qy 121 YTLGADLTPEGEMVNSGKIFQYVESUTRNPIIRSGSGSRVJASGKKFIEGQST 180
Db 120 YTLGADLTPEGEOQLVNSGKIFQYVESUTRNPIIRSGSGSRVJASGKKFIEGQST 179
Qy 181 KLADPGSQHPASPVINIIPEGSGYNNITLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 180 KLADPGCA-TNRAAPALSIPTSESTENNTLDHGVCTKFEASQLGDEVAANFTALFAPDIR 238
Qy 241 ARLEADLPGLVTLDEDDVYVLMDCPFDTVARTSDATELSFPCALFTHDEWIQDYLOSIG 300
Db 239 ARAEKLHPGVTLDLDEDDVYVLMDCSEDTVARTSDATELSFPCALFTHDEWIQDYLOSIG 298
Qy 301 KYGAGNPLGAGGCGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 299 KYGAGNPLGAGGCGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 358
Qy 361 HDNTMISIFFALGNGTKPLSTTSVESIEETDGYASWSVTPFAARAYVEMMOCAKEP 420
Db 359 HDNSMWSIFFALGNGTKPLSTTSVESIEETDGYASWSVTPFAARAYVEMMOCAKEP 418
Qy 421 LVRVLNDRVPLHGCADVCKLGRCKRDFVEGLSFARSGGNWAECEFA 467
Db 419 LVRALINDRVPLHGCADVCKLGRCKRDFVEGLSFARSGGNWAECEFA 465

RESULT 3
ID 093838 PRELIMINARY; PRT; 467 AA.
AC 093838;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PHYTASE.
GN PHA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-57;
RA Nagashima T.; Kondo H.; Anazawa H.; Terasaki Y.;
RT "Phytase having high-affinity for phytic acid";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022700; CAB19824.1; -.
DR HSP; P34752; 1IHP.
DR InterPro: IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR


```

FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 466 AA; 51055 MW; F2AECECIAF7C22C4 CRC64;

Query Match 74.7%; Score 1839; DB 3; Length 466;
Best Local Similarity 73.0%; Pred. No. 2.8e-139;
Matches 340; Conservative 49; Mismatches 77; Indels 0; Gaps 0;

Qy 1 MGFFVLLSIATLFGSTGCTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60
Db 1 MGFFVLLSIATLFGSTGCTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60

Qy 61 SAISPDVDDCVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFLKTYN 120
Db 61 SPFPLDVPDCHITFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFLKTYN 120

Qy 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVITASAETIEGFQSA 180
Db 121 YSMGSENLNPFGRNQLQDLGAQFYRYRYKALARKIVPFIIRASGSDRVITASAETIEGFQSA 180

Qy 181 KLADPGSQHQAQSPVINVIIPEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 ROGDDHANPHQSPRVVDVVAIPEGSAIYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240

Qy 241 ARLEADLPVGLTDEDVYVLMDCMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
Db 241 KRLADLPVGLTDEDVYVLMDCMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300

Qy 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATFPLNATLYADFS 360

Qy 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420
Db 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420

Qy 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDFFVEGLSFARSGGNWAECE 466
Db 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDFFVEGLSFARSGGNWAECE 466

RESULT 6
000085 ID 000085 PRELIMINARY; PRT; 466 AA.
AC 000085;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE, 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHA.
OS Aspergillus terreus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
ON NCBI_TaxID=33178;
RX [1]
RC STRAIN=9A1;
RX MEDLINE=97177792; PubMed=90252398;
RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
RT "The phytase subfamily of histidine acid phosphatases: isolation of
RT genes for two novel phytases from the fungi Aspergillus terreus and
RT Myceliophthora thermophila."
RL Microbiology 143:245-252(1997).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHITATE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O -> D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
```

```

CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
CC 2.5 TO 7.5, WITH MAXIMAL ACTIVITY AT PH 5.5. ALSO ACCEPT 4-
CC NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO
CC MORE ACIDIC PH VALUES.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL: U59805; AAB52507.1; .
DR HSSP: P34752; LIHP.
DR InterPro: IPR000560; .
DR Pfam: PF00328; acid_phosphat; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 466 REQUIRED FOR BINDING SUBSTRATE (BY
FT ACT_SITE 82 82 SIMILARITY).
FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
FT DISULFID 31 40 BY SIMILARITY.
FT DISULFID 71 414 BY SIMILARITY.
FT DISULFID 215 465 BY SIMILARITY.
FT DISULFID 264 282 BY SIMILARITY.
FT DISULFID 436 444 BY SIMILARITY.
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 466 AA; 51093 MW; 21DCB559C96AB66 CRC64;

Query Match 71.5%; Score 1761; DB 3; Length 466;
Best Local Similarity 69.5%; Pred. No. 5.1e-133;
Matches 324; Conservative 58; Mismatches 84; Indels 0; Gaps 0;

Qy 1 MGFFVLLSIATLFGSTGCTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60
Db 1 MGFLAIVLSVALLFRSTSGTGLPGRKHSCDVTGGYQCPEISHLWGTSPYFSLADE 60

Qy 61 SAISPDVDDCVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFLKTYN 120
Db 61 SPFPLDVPDCHITFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFLKTYN 120

Qy 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVITASAETIEGFQSA 180
Db 121 YSLDSEELTPFGRNQLRDLGAQFYRYRYKALARKIVPFIIRASGSDRVITASAETIEGFQSA 180

Qy 181 KLADPGSQHQAQSPVINVIIPEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 RQDDHANPHQSPRVVDVVAIPEGSAIYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240

Qy 241 ARLEADLPVGLTDEDVYVLMDCMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
Db 241 QRLADLPVGLTDEDVYVLMDCMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300

Qy 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATFPLNATLYADFS 360

Qy 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420
Db 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420

Qy 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDFFVEGLSFARSGGNWAECE 466
Db 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDFFVEGLSFARSGGNWAECE 466

RESULT 7
000096 ID 000096 PRELIMINARY; PRT; 466 AA.
```

000096;
 01-JUL-1997 (Tremblrel. 04, Created)
 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 01-NAR-2001 (Tremblrel. 16, Last annotation update)
 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAPHOSPHATE
 PHA).
 OS Talaromyces thermophilus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Talaromyces.
 OX NCBI_TaxID=28565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20186;
 RX MEDLINE=98007872; PubMed=9349716;
 RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,
 RA van Loon A.P.G.M.;
 RT "Cloning of the phycases from *Emicella nidulans* and the thermophilic
 fungus *Talaromyces thermophilus*."
 RL Biochim. Biophys. Acta 1353:217-223(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- MISCELLANEOUS: HAS POTENTIAL AS FOOD AND FEED ADDITIVE. IT CAN
 FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
 USED AS FOOD FOR MONOGASTRIC ANIMALS.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U59802; AB96873.1; -;
 DR HSP; P34752; LIHP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 466
 FT ACT_SITE 79 79 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 357 357 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 28 37 BY SIMILARITY.
 FT DISULFID 68 410 BY SIMILARITY.
 FT DISULFID 212 461 BY SIMILARITY.
 FT DISULFID 261 278 BY SIMILARITY.
 FT DISULFID 432 440 BY SIMILARITY.
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51450 MW; FC4575B521A5C929 CRC64;
 Query Match 69.8%; Score 1719; DB 3; Length 466;
 Best Local Similarity 69.3%; Pred. No. 1.2e-129;
 Matches 325; Conservative 42; Mismatches 94; Indels 8; Gaps 3;
 QY 1 MGFWVLLS--IATFGSTGCTALGPRGNHSCDVTGGYQCPEISHLWGTSPYFSLA 58
 Db 1 MSLLLLVSLGLVALYVSRN-----PHVDSHSCNTVEGGYQCRPEISHSHSQSPFSLA 55
 QY 59 DESAISPDVDDCVTFVQVLSRHGARYPTSSKASKYASALIEAIQKNATAFKGYAFPLKT 118
 Db 56 DOSEISPDVQNCITVFQVLSRHGARYPTSSKASKYASALIEAIQKNATAFKGYAFPLKT 115
 QY 119 NYNTLGADLTPFGENQWNSGKIFRYRYKALARKIVPFTIRASGSDRVVASAEKFEIQQ 178
 Db 116 YRYLGANDLTPFGENQWNSGKIFRYRYKALARKIVPFTIRASGSDRVVASAEKFEIQQ 175
 QY 179 SAKLADPGSQPHQASPVNVIIEGSGYNNLTLDGTCTAFEDSELGDDVEANFTALFAPA 238
 Db 176 SAKYLDPHSKDHPAPTINVIIIEGSGYNNLTLDGTCTAFEDSELGDDVEANFTALFAPA 235

QY 239 IRARLEADLPGVTLTDEDDVYVLMDCPPFTVARTSDATSELSFPCALFTHDWIOYDLOS 298
 Db 236 ILEKIKHLPVGLAVSDVPLMDLCPFTLARNHTDT-LSPFCALSTOEWQAYDYOQS 294
 QY 299 LGKYGYGAGNPLGPAQGVGFANELLARLTHSPVQDHTSTNHTLDSNPATFPLNATLYAD 358
 Db 295 LGKYGGNGGNPLGPAQGVGFANELLARLTHSPVQDHTSTNHTLDSNPATFPLNATLYAD 354
 QY 359 FSHDNTMISTIFFALGYNGTKPLSTTSVESIETDGYASWTVPFAARAYVEMQCAEK 418
 Db 355 FSHDNTMISTIFAALGYNGTKPLSTTSVESIETDGYASWTVPFAARAYVEMQCAEK 414
 QY 419 EPLRVLVNDRVPLHCCAVDKLGRCKRDDFVEGLSFARSGGNWACFA 467
 Db 415 EPLRVLVNDRVPLHCCAVDKLGRCKRDDFVEGLSFARSGGNWACFA 463
 RESULT 8
 ID 000107 PRELIMINARY; PRT; 487 AA.
 AC 000107;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DE 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 PHA).
 GN Thielavia heterothallica.
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Chaetomiaceae; Thielavia.
 OX NCBI_TaxID=78579;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97177792; PubMed=9025259;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,
 RA van Loon A.P.G.M.;
 RT "The phytase subfamily of histidine acid phosphatases: isolation of
 genes for two novel phytases from the fungi *Aspergillus terreus* and
Myceliophthora thermophila."
 RL Microbiology 143:245-252(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO
 ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA
 SHIFTED TO MORE ACIDIC PH VALUES.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U59806; AB52508.1; -;
 DR HSP; P34752; LIHP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 75 POTENTIAL.
 FT CHAIN 1 487
 FT DOMAIN 267 270 POLY-SER.
 FT DOMAIN 423 433 POLY-GLY.
 FT ACT_SITE 75 75
 FT ACT_SITE 76 76
 FT ACT_SITE 368 368
 FT DISULFID 26 35
 FT DISULFID 64 421 BY SIMILARITY.
 FT DISULFID 208 485 BY SIMILARITY.
 FT DISULFID 260 289 BY SIMILARITY.
 FT DISULFID 456 464 BY SIMILARITY.
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).

RC TISSUE=LIVER;
 RX MEDLINE=98028656; PubMed=9359836;
 RA Craxton A., Caffrey J.J., Burkhardt W., Safrany S.T., Shears S.B.;
 RT "Molecular cloning and expression of a rat hepatic multiple inositol
 RL polyphosphate phosphatase";
 RN Biochem. J. 328:75-81(1997).
 RN [2]
 RX CATALYTIC ACTIVITY.
 RP MEDLINE=91358435; PubMed=1653239;
 RA Nogimori K., Hughes P.J., Glennon M.C., Hodgson M.E., Putney J.W. Jr.,
 RA Shears S.B.;
 RT "Purification of an inositol (1,3,4,5)-tetrakisphosphate 3-phosphatase
 RT activity from rat liver and the evaluation of its substrate
 RT specificity";
 RL J. Biol. Chem. 266:16499-16506(1991).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INOSITOL 1,3,4,5-
 CC TETRAPHOSPHATE TO INOSITOL 1,4,5-TRIPHOSPHATE AND INOSITOL
 CC 1,3,4,5,6-PENTAKISPHOSPHATE TO INOSITOL 1,4,5,6-TETRAKISPHOSPHATE
 CC AND THEN TO INOSITOL 1,4,5-TRISPHOSPHATE. DEPHOSPHORYLATES
 CC INOSITOL HEXAKISPHOSPHATE WITHOUT SPECIFICITY TOWARDS A PARTICULAR
 CC PHOSPHATE GROUP.
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY AND LIVER.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL, AF012714; AAC53453.1; -
 DR InterPro; IPR000560; -
 DR InterPro; IPR000886; -
 DR Pfam; PF00328; acid phosphat; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; FALSE_NEG.
 KW Hydrolase; Glycoprotein; Phosphorylation; Multifunctional enzyme.
 FT ACT_SITE 59
 FT MOD_RES 65
 FT MOD_RES 146
 FT MOD_RES 176
 FT MOD_RES 201
 FT MOD_RES 218
 FT MOD_RES 391
 FT MOD_RES 412
 FT MOD_RES 435
 FT MOD_RES 447
 FT MOD_RES 447
 FT CARBOHYD 206
 FT CARBOHYD 445
 FT VARIANT 3
 FT CONFLICT 327
 FT SITE 448
 FT SEQUENCE 451 AA; 51592 MW; EBI05512A03020B CRC64;

Query Match 7.7%; Score 189; DB 11; Length 451;
 Best Local Similarity 20.7%; Pred. No. 5.5e-07;
 Matches 89; Conservative 66; Mismatches 193; Indels 82; Gaps 16;
 QY 52 SPYFS-----LAESAISPD---VPDCRVTFVQVLSRHGARYPTSSASKAYS 96
 DB 14 SPYFGTKTRYEDVNPWLLGDPVAPRRDPPELLAGTCTPVQLVALIRHGTTRYPTTKQIRKLR 73
 QY 97 ALIEAIOKNATFAGKAFKLTNTYTLGADLLPFGENQMVNSGKIFRYRYKALARKIVP 156
 DB 74 QLOGLLQTRSVGDSGRVAALDQWPLWYDD---WMDGOLVERGQDMQRLALRLAALFP 130
 QY 157 --FIR-----AGSDRVIAAEKFIQFQSAKLADPGSQPHQAS-----APAIRARLE-----PVINVI 199
 DB 131 DLFCRENGRLRLITSSKRCVSSAAFLQGLW--QHYPGLPPDPVSDMECPPRVN-- 186
 QY 200 IPEGSGYNTLDTGCTAFEDSEIGDDVEANFTALF-----APAIRARLE-----ADL 247
 DB 187 ----DKLMRFEDH--CEKEL-----TEVERNATLYHVEAFKTPGEMQTVLKKVAATLQV 235
 QY 248 PGVTLTDEDVYVLMDCMPPFTVARTSDATSELPFCALFTHDEWIQYDYLQSLGKYGYGA 307
 DB 236 PVNNLNADLIQVAFFTCGSDIAIQGVH-----SPWCDVDFDQVDDAKVLEYLNDLKOYWKRSY 291

QY 308 GNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMIS 367
 DB 292 GYAINSRSSCNLFQDIFLHLDKAVEQKORSQ-----PVSSVILQFGHAETLLP 340
 QY 368 IFFALGLYNTKPLSTITSVESIEETDGYSASWTVPFAARAYVEMMOCAQKEP---LVR 423
 DB 341 LLSIMGYFKDKPLTAYNEEQVHRE-FRSGHIVPASNLIIFVLYHCEDAQTPQEFQIQ 399
 QY 424 VLVNDRVVPL 433
 DB 400 MLLNEKVLP 409
 RESULT 15
 Q922L6 PRELIMINARY; PRT: 481 AA.
 ID Q922L6
 AC Q922L6;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
 GN MINPPI OR MIPP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046508; AAD02434.1; -
 DR MGI; MGI:1336159; Minppl.
 DR InterPro; IPR000560; -
 DR InterPro; IPR000886; -
 DR Pfam; PF00328; acid phosphat; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
 SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;
 Query Match 7.6%; Score 186; DB 11; Length 481;
 Best Local Similarity 21.0%; Pred. No. 1e-06;
 Matches 99; Conservative 66; Mismatches 206; Indels 100; Gaps 18;
 QY 11 ATLFGSTGALGPRNSHSCDVTVDGYQCPEISHLWGTSPYFS-----IA 58
 DB 21 AALLSSFARCSLPGRGD-----PVASVL-----SPYFGTKTRYEDANPWLIV 62
 QY 59 DESAISPD---VPDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATFAGKAYF 115
 DB 63 DPVAPRRDPPELLAGTCTPVQLVALIRHGTTRYPTTKQIRKQLQGLLQTRSDGSGQA 122
 QY 116 LKTYNTLGCADLLTPGENQMVNSGKIFRYRYKALARKIVP--FIR-----AGSD 164
 DB 123 AALAEPLWYGD---WMDGOLVERGQDMQRLALRLAALFPDLFSRENYDRRLITSSKH 179
 QY 165 RVIAAEKFIQFQSAKLADPGSQPHQAS-----PVINVIPEGSGYNTLDTGCTAF 218
 DB 180 RCVDSAAFLQGLW--QHYPGLPPDPVSDMECPPRIN-----DKLMRFEDH--CEK 229
 QY 219 EDESEIGDDVEANFTALF-----APAIRARLE-----ADLPGVTLTDEDVYVLMDCMP 266
 DB 230 L-----TDVERNATLYHVEAFKTPGEMQVKLVKVAATLQVPMNSLNADLIQVAFFTCF 284
 QY 267 DTVARTSDATSELPFCALFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFANELIAR 326
 DB 285 DLAIKGVH-----SPWCDVDFDQVDDAKVLEYLNDLKOYWKRSYGYTINSSSCNLFIDFLH 340
 QY 327 LKHSVPQDHTSTNHTLDSNPATPLNATLYADFSHNTMISIFFALGLYNGTKPLSTTSV 386
 DB 341 LDKAVEQKORSQ-----PVSSVILQFGHAETLLPILLSLMGYFKQKEPLTAYNF 389

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:28 ; Search time 48.78 Seconds
(without alignments)
729.264 Million cell updates/sec

Title: US-09-488-265-29

Perfect score: 2462

Sequence: 1 MGVFVLLSIATLFGSTSGT.....DFVEGLSFARSGGNWAECA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	76.3	467	1 JN0889	3-phytase (EC 3.1.1)
2	1875	76.2	467	1 JN0656	3-phytase (EC 3.1.1)
3	1814	73.7	441	1 JN0482	3-phytase (EC 3.1.1)
4	395	16.0	467	2 S53476	acid phosphatase (EC 3.1.1)
5	395	16.0	467	2 S48996	acid phosphatase (EC 3.1.1)
6	392	15.9	467	1 PABYC	acid phosphatase (EC 3.1.1)
7	380	15.4	467	1 PABYC	acid phosphatase (EC 3.1.1)
8	355	14.4	468	2 JC4285	acid phosphatase (EC 3.1.1)
9	354	14.4	468	2 S52495	acid phosphatase (EC 3.1.1)
10	347.5	14.1	479	1 JN0890	acid phosphatase (EC 3.1.1)
11	346.5	14.1	479	1 JN0890	acid phosphatase (EC 3.1.1)
12	297.5	12.1	463	2 T39929	thiamin-repressibl
13	297	12.1	453	1 A25326	acid phosphatase (EC 3.1.1)
14	278.5	11.3	463	2 S14119	acid phosphatase (EC 3.1.1)
15	149	6.1	465	2 JE0369	histidine acid pho
16	145.5	5.9	468	2 A86233	hypothetical prote
17	130	5.3	755	2 T19118	acid phosphatase h
18	124.5	5.1	421	2 S14742	acid phosphatase h
19	119	4.8	449	2 T15933	hypothetical prote
20	114.5	4.7	423	2 A33395	acid phosphatase (EC 3.1.1)
21	112	4.5	721	2 T27570	hypothetical prote
22	106.5	4.3	380	2 T16883	hypothetical prote
23	104	4.2	354	2 T21241	hypothetical prote
24	103	4.2	1225	2 T16346	hypothetical prote
25	102.5	4.2	423	1 S06167	hypothetical prote
26	102	4.1	452	2 T20556	acid phosphatase (EC 3.1.1)
27	101.5	4.1	680	2 S63698	hypothetical prote
28	101	4.1	563	2 S33957	NADPH--ferrhemopr
29	100.5	4.1	730	2 T16455	coat protein gamma
					hypothetical prote

30 98.5 4.0 381 2 JH0152 acid phosphatase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
31 98.5 4.0 642 2 E69144 probable formate C
32 98.5 4.0 680 2 S63895 NADPH--ferrhemopr
33 97.5 4.0 715 4 TVMSMY transforming prote
34 96.5 3.9 275 2 S09774 hypothetical prote
35 96.5 3.9 386 1 JH0610 acid phosphatase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
36 96.5 3.9 1433 2 A71444 acid phosphatase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
37 96 3.9 413 2 T18945 glycine dehydrogen
38 95.5 3.9 473 2 G84312 NADPH--ferrhemopr
39 95.5 3.9 680 2 A37890 hypothetical prote
40 95.5 3.9 845 2 T17291 lysine-specific cy
41 95.5 3.9 1732 2 T30836 probable peptide s
42 95.5 3.9 4924 2 T50176 alpha-amylase (EC 3.2.1.3) A precursor - human
43 95 3.9 605 2 JH0638 UL25 protein - hum
44 94.5 3.8 580 1 WMBEWS exo-poly-alpha-gal
45 94.5 3.8 602 2 A36715

ALIGNMENTS

RESULT 1
JN0889
3-phytase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
C:Species: Aspergillus awamori
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0889
R:PIDdington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0889
A:Molecule type: DNA
A:Residues: 1-467 <PID>
A:Cross-references: GB:I02421; NID:q166518; PIDN:AAAL6898.1; PID:g166519
A:Experimental source: strain ALK0243
A:Note: part of the sequence, including the amino end of the mature protein, was conf
C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic
C:Genetics:
A:Gene: phyA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-467/Product: 3-phytase A #status experimental <MAT>
F:27.59,105,120,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 76.3%; Score 1879; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 1.4e-143;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTSGTALPGRNHSCDVTDDGGVQCQFPETSHLWGTVPYFSLADE 60
DB 1 MGVSALLPLLYLLAGVTSGTAVASRNQSTCDTVDQGYQCFSETSHLWGVAPFFSLANE 60
QY 61 SAISPDPVDDCRVTQVLSRHGARYPTSSKAYSALEIAIQKNATKGYAFLKTYN 120
DB 61 SAISPDPVAGCRVTFAQVLSRHGARYPTESKGYSALEIQNNVTTFDKYAFKTYN 120
QY 121 YTLGADDTLPGENOMVNGIKFYRRYKALARKIVPFIRASGSDRVIASAEKFTGFQSA 180
DB 121 YSLGADDTLPGEQELVNSGKIFQYRSLRNIPFIRSSGSRVITASGEKFTGFQST 180
QY 181 KLADPGSOPHOASPVINIIPEGSGYNNLTDLHGTCTAFEDSELGDDVEANFTALFAIR 240
DB 181 KLKDPAGPGGSSKIDVVIIEASSSNNLTDPGCTCTFEDELADTVEANFTATFASIR 240
QY 241 ARLEADLVGTLTDEVDVYLMDCPFDTFVARTSDATSLSPFCALFTHDEWIOYDLOSGL 300
DB 241 ARLEADLVGTLTDEVDVYLMDCPFDTFVARTSDATSLSPFCALFTHDEWIOYDLOSGL 300

Db 241 ORLENDLSGVLTDTTEVYLLMDMCSFDITSTSTVDTKLSPDCDLTFHDEWIHYDLSQK 300
 QY 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYIGHGAGNPLGPTQGVGYNELIARLTHSPVHDTSSNHTLDSNPATFPLNATLYADFS 360
 QY 361 HDNTMISIFALGLYNGTKPLSTSVESIEETDGYASAWTPFAARAYVEMMQCAQEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENTITQDGFSSAWTPFAARLYVEMMQCAQEP 420
 QY 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 Db 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 RESULT 2
 JN0656
 3-phytase (EC 3.1.3.8) A precursor - Aspergillus niger
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phvA protein
 C:Species: Aspergillus niger
 C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
 R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.
 A.M.J.J.
 Gene 127, 87-94, 1993
 A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phvA)
 A:Reference number: JN0656; MUID:93252284
 A:Accession: JN0656
 A:Molecule type: DNA
 A:Residues: 1-467 <N>
 A:Cross-references: GB:216414; NID:q2392; PIDN:CAA78904.1; PID:g2393
 A:Experimental source: strain NRRL3135
 A:Note: Parts of the sequence, including the amino end of the mature protein, were confi
 C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and ino
 C:Genetics:
 A:Gene: phvA
 A:Introns: 15/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phos
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-467/Product: 3-phytase A #status experimental <MA>
 F:27,59,105,120,207,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #sta
 F:81,361/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted
 Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 3e-143;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;
 QY 1 MGVEVVLSTATLFGSTSGTALGPRGNHSCDFTDGGYOCFPEISHLWGTSPYFSLADE 60
 Db 1 MGVSALLPYLLSGVTSGLAVPASRNQSCDFTDGGYOCFSETSHLWQYAPFFSLANE 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKYASALIEATOKNATAPKGYAFELKTYN 120
 Db 61 SVISPEVPACRVTFQAQLSLRHGARYPTDSKGYASALIEEQONATTFDGYAFELKTYN 120
 QY 121 YTLGADDLTPFGENOMVNSGKIKYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 Db 121 YSLGADDLTPFGQELVNSGKIKYRYKALARKIVPFIASGSDRVIASAEKFIQFQST 180
 QY 181 KLADPGSOPHOAGSPVNIIVPEGSYNNNTLDHGTCTAFEDSELDGDDVEANFTALFAPAIR 240
 Db 181 KLKDPRAQPGQSPKIDVLISEASSNNTLDPCCTVFEDSELADTVEANFTAFVPSIR 240
 QY 241 ARLEADLPVTLTDEDVYVLMDCPFTVARTSDATSELPFCALFTHDEWIQYDYLQSLG 300
 Db 241 ORLENDLSGVLTDTTEVYLLMDMCSFDITSTSTVDTKLSPDCDLTFHDEWIHYDLSQK 300
 QY 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYIGHGAGNPLGPTQGVGYNELIARLTHSPVHDTSSNHTLDSNPATFPLNATLYADFS 360
 QY 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 Db 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467

QY 361 HDNTMISIFALGLYNGTKPLSTSVESIEETDGYASAWTPFAARAYVEMMQCAQEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENTITQDGFSSAWTPFAARLYVEMMQCAQEP 420
 QY 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 Db 421 LVRVLNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 RESULT 3
 JN0482
 3-phytase (EC 3.1.3.8) A - Aspergillus ficum
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phvA protein
 C:Species: Aspergillus ficum
 C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
 R:Ullah, A.H.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 192, 747-753, 1993
 A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemi
 A:Reference number: JN0482; MUID:93249451
 A:Accession: JN0482
 A:Molecule type: protein
 A:Residues: 1-441
 A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
 R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 178, 45-53, 1991
 A:Title: Cyclohexanedione modification of arginine at the active site of Aspergillus
 A:Reference number: PN0023; MUID:91298982
 A:Accession: PN0023
 A:Molecule type: protein
 A:Residues: 48-70
 C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
 F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydrate (Asn) (covalent) #sta
 F:58,338/Active site: Arg, His #status predicted
 F:59/Active site: His (phosphohistidine intermediate) #status predicted
 Query Match 73.7%; Score 1814; DB 1; Length 441;
 Best Local Similarity 77.2%; Pred. No. 2.3e-138;
 Matches 338; Conservative 39; Mismatches 61; Indels 0; Gaps 0;
 QY 27 NSHSCDFTVDDGGYOCFPEISHLWGTSPYFSLADESAISPDVDDCRVTFVQVLSRHGARY 86
 Db 4 NQSSCDFTVDDGGYOCFSETSHLWQYAPFFSLANESVISEPVEPAGCRVTFQAQLSLRHGARY 63
 QY 87 PTSASAKYASALIEATOKNATAPKGYAFELKTYNLTGADDLTPFGENOMVNSGKIKFYR 146
 Db 64 PTDKGGKYSALIEEQONATTFDGYAFELKTYNLSLUGADDLTPFGQELVNSGKIKFYR 123
 QY 147 YKALARKIVPFIASGSDRVIASAEKFIQFQSAKLADPGSOPHOAGSPVNIIVPEGSY 206
 Db 124 YESLTRLNVPFIASGSDRVIASCKKFIQFQSTKLKDPRAQPGQSPKIDVLISEASS 183
 QY 207 NNTLDHGTCTAFEDSELDGDDVEANFTALFAPAIRLEADLPVTLTDEDVYVLMDCPF 266
 Db 184 NNTLDPGCTCTFEDSELADTVEANFTAFVPSIRQRIENLDSGVTLTDTTEVYLMDCSF 243
 QY 267 DTVARTSDATSELPFCALFTHDEWIQYDYLQSLGKYYGYGAGNPLGPAQGVAFANELIAR 326
 Db 244 DTISTSTVDTKLSPDCDLTFHDEWIQYDYLQSLGKYYGYGAGNPLGPAQGVAFANELIAR 303
 QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMISIFALGLYNGTKPLSTTV 386
 Db 304 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNGIISILFALGLYNGTKPLSTTV 363
 QY 387 RSIETDGYASAWTPFAARAYVEMMQCAQEPFLVRVLNDRVPLHGCVDKLGCRK 446
 Db 364 ENITDGFSSAWTPFAARLYVEMMQCAQEPFLVRVLNDRVPLHGCVDKLGCRK 423
 QY 447 DDFVEGLSFARSGGNWAE 464

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R:Feldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45927
A:Accession: S45960
A:Molecule type: DNA
A:Residues: 1-467 <FE2>
A:Cross-references: EMBL:Z35961; NID:G536362; PIDN:CAA85045.1; PID:G536363; GSPDB:GN0000
R:Baywa, W.; Mayhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
Nucleic Acids Res. 12, 7721-7739, 1984
A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeast
A:Reference number: S05794; MUID:85037940
A:Accession: S05794
A:Molecule type: DNA
A:Residues: 1-218, 'MKT', 222-467 <BAJ1>
A:Cross-references: EMBL:X01080; NID:G4148; PIDN:CAA25557.1; PID:G758281
R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G.
Mol. Cell. Biol. 6, 1855-1865, 1986
A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
A:Reference number: A93074; MUID:87064474
A:Accession: A25241
A:Molecule type: DNA
A:Residues: 1-44 <TAI>
C:Genetics:
A:Gene: SGD:PHO3; MIPS:YBR092c
A:Cross-references: SGD:S0000296; MIPS:YBR092c
A:Map position: 2R
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-467/Product: acid phosphatase constitutive #status predicted <MAT>
F:75/Active site: His (phosphohistidine intermediate) #status predicted
F:97/103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (covalent)
F:97/Active site: His #status predicted

Query Match 15.4%; Score 380; DB 1; Length 467;
Best Local Similarity 26.3%; Pred. No. 9.1e-23;
Matches 114; Conservative 61; Mismatches 196; Indels 62; Gaps 11;
QY 49 GTPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIBAIQKATA 108
Db 42 GGAGYFSPFGDYGISRLDPECCENKQLQMLARHGERTYPSKATIMTKYKLSNYTRQ 101
QY 109 FKGKTAFLK-TNYTLGADD-----LTPF-GENOMVNSGKIFRYRYKALAR 152
Db 102 FNGSLFSLNDDEYFFIRDDDLLEMETTFANSQNLVNLPTGEMDAKRHRAREFLAQYGYMPE 161
QY 153 KIVPF-IRASGSDRVASAEKIEFGFSQAKLADPGSQHOASPVINVIPEGSGYNTLD 211
Db 162 NOTSPFPAASSERVHDIAQYFIDGL-----GDQFNISLOTSEAMSAGA---NTLS 210
QY 212 HGCTAFEDSELGDDVEANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCPFDTVAR 271
Db 211 AGNACPGWEDANDDILDKYDTYLLDDIAKRLNKENKGLNLSKANTLFANCAVELNAR 270
QY 272 TSDATLSPFCALFTHDEWIQDYLSLQSLGKYYGAGNPLGPAQGVGFANELIARLTHSP 331
Db 271 -----GYSDVCDIFTEDELVRYSYQDLSVSYQDQPGYDMIRSVGANLFPNATLKLKQSE 325
QY 332 VODHSTNHTLDSNPATPLNATLYADFSHDNTMISIFFALGLYNGTKPLSTTSVESIEE 391
Db 326 TQD-----LKVWLSFHDITDILNLTATGIIDDKNNLTAEYVFFMGN 367
QY 392 TDGYSASWTVPFAARAYVEMMQCQAEKEPLRVLVNDRVPLVHGCAVDKLGCKRDDF-- 449
Db 368 T---FHKSWVPOGARVYTEKFC--SNDTYVRYVINDAVPIETCTSGTGSCFCEINDFYD 423
QY 450 -----VEGLSFAR 457
Db 424 YAEKRVAGTDFLK 436

RESULT 8
JC4285
N:Alternate names: orthophosphoric-monoester phosphohydrolase (acid optimum); Pho
C:Species: Pichia pastoris
C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999
C:Accession: JC4285
R:Payne, W.E.; Cannon, P.M.; Kaiser, C.A.
Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast Pichia pastoris: Characterization
A:Reference number: JC4285; MUID:96001238
A:Accession: JC4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:G881955; PIDN:AAA85503.1; PID:G881956
A:Experimental source: GS115
C:Genetics:
A:Gene: pho1
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:84/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:345/Active site: His #status predicted

Query Match 14.4%; Score 355; DB 2; Length 468;
Best Local Similarity 25.4%; Pred. No. 9.4e-21;
Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;
QY 6 VLLSTATL---FGSTGATLGRPGNSHSCDVTDDGGVQCQPFESHLMGTSPY--FSLADE 60
Db 10 IIALATLOSFAVELQHVLRVNDPRYPQRT-DDQYNILRHGLGL-GPYIGYNGWGIAAE 67
QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIE-AIQKNATAFKKAYALKY 119
Db 68 SEI-----ESCIDQHLMLRHGERYPSTNVGKQLEALYQKLLDADVEVTPGLSFFQY 122
QY 120 NYTLG-----ADLTP---FGENQMVNSGKIFRYRYKAL-----ARKIVPFIASGSDR 165
Db 123 DYFVSDAAWVEQETTKGFYSGLNTAFDGTTLRERYDHLINTSEEGKL--SWAGSQER 180
QY 166 VTASAEKFTIEGFSQAKLADPGSQHOASPVINVIPE-----GSGYNTLDHGCTCT-----A 217
Db 181 VYDTAKYFAQGFMSNYTD-----MVEYVALEEEKSQGLNSLTARISCPNYNSHI 230
QY 218 FDSSELGDDV---EANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCPFDTVARTSD 274
Db 231 YKDGDFPDIAAREAD-----RLNTLSPGNITADDTPTIYALCYGGEINVRGE-- 278
QY 275 ATLSLSPFCALFTHDEWIQDYLSLQSLGKYYGAGNPLGPAQGVGFANELIARLTHSPVD 334
Db 279 ----SFCDVLSREALYLYALRDLGWYVNGVNGNPLGKTIGYVYAN----- 321
QY 335 HTSTNHTLDSNPATPLNATLYADFSHDNTMISIFFALGLYNGTK-PLSTTSVESIEEFD 393
Db 322 --ATROLLENTEAD-PRDYPYLFYSFSDHDLQVFTSLGLFNVTDLPL-----DQIQFOT 373
QY 394 GYSASWTVPFAARAYVEMMQC--QAEKEPLRVLVNDRVPLVHGCAVDKLGCKRDDFVE 451
Db 374 SPKSTEIVPMGARLLTERLLCTVEGEEKYVVRTILNDVFPPLSDCSSGPGFSCPLNDIYS 433
QY 452 GLSFARSGGNWAE 464
Db 434 RLEALNEDSDFAE 446
RESULT 9
S52495
N:Alternate names: acid phosphatase homolog YDL024c - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D2815

A: Molecule type: protein
A: Residues: 65-66, 68-93 <ULL>
C: Comment: This enzyme, previously characterized as an acid phosphatase (EC 3.1.3.2), has
C: Genes: phbB
A: Gene: phbB
A: Introns: 261/1; 300/2; 335/2
C: Superfamily: yeast acid phosphatase
C: Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
F: 1-19/Domain: signal sequence #status predicted <SIG>
F: 20-479/Product: 3-phytase #status experimental <NAT>
F: 81,337/Active site: Arg, His #status predicted
F: 82/Active site: His (phosphohistidine intermediate) #status predicted
F: 106,191,227,250,315,340,425,442,458/Binding site: carbohydrate (Asn) (covalent) #status
Query Match 14.1%; Score 346.5; DB 1; Length 479;
Best Local Similarity 25.3%; Pred. No. 4.7e-20;
Matches 119; Conservative 54; Mismatches 170; Indels 63; Gaps 17;
Qy 49 GTSPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAY-SALTEATOKNAT 107
Db 53 GPYSERVSY-----GIADPTTCCEVDQVIMVYKRGERYPSAGKSIEALAKVYSINT 108
Qy 108 AFKGYAFKTYNYTL-----GADLT-PP-GENQMVNSGKIFRYKAL--ARKIYVPF 157
Db 109 EYKGLAFPLNDWTYVYPNCEYNAETTSQVAGLLDAYNHGNDYKARYGHLWNGETVVPF 168
Qy 158 IRASGRDVIASAEKFIQFOSAKLADPGSQPHQASPVINII-PEGSYNTLDHGCT 216
Db 169 F-SSGVRVIETARKFGGFGYN-----YSTNAALNIISEVNGADSLP--TCD 217
Qy 217 AFESELDGDDVEANFTALFAIPARLEADLPVTLTDEDVYVLMDCPFDVTARTSDAT 276
Db 218 TDNDQTTCDNLTYQLPQFKVAA--ARLNSQNPGMNLTASDVYVLMVASFELNAR----- 270
Qy 277 ELSPPCALFTDHWITQYDYLQSLGKYGYGAGNPLGPAGQGVGFANELLARLTHSPVDHT 336
Db 271 PFSNMINAFTQDQWVSFGYVEDLNYVCAGPGDKNMAAGVAVYANASLTLLNQGP----- 325
Qy 337 STNHTLDSNPATFPLNATLYADFSHDNTMISIFFALGLY--NGTKPLSTTSVESTETD 394
Db 326 -----KEAGP-----LFTLAHDNTITPILAAGLVLIPIEDLPL-----DRVAFGNP 367
Qy 395 YSASWTVPFAARAYVEMMQCOA-----EKEPLRVRLVNDRVVPLHGC 436
Db 368 YSIGNIVPMGGHLTIERSQATLSQATLSQATLSQATLSQATLSQATLSQATLSQATLSQ 413
RESULT 12
T39929
thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)
C: Species: Schizosaccharomyces pombe
C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C: Accession: T39929
R: Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.
submitted to the EMBL Data Library, May 1998
A: Reference number: 221857
A: Accession: T39929
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-463 <LYN>
A: Cross-references: EMBL:AL023286; PIDN:CAA18863.1; GSPDB:GN00067; SPDB:SPBC21H7.03C
A: Experimental source: strain 972h; cosmid c21H7
C: Genes: phbB
A: Gene: SPDB:SPBC21H7.03C
A: Map position: 2
C: Superfamily: yeast acid phosphatase

Query Match 12.1%; Score 297.5; DB 2; Length 463;
Best Local Similarity 25.3%; Pred. No. 4e-16;
Matches 114; Conservative 67; Mismatches 203; Indels 67; Gaps 15;

Qy 46 HLWGTYS-----PYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTS--SASKAYSA-- 97
Db 35 HL-GTLSYPHYPFDGLDSA-----FPTCEIQVHLLQRHGSRNPTGDVTATDVYSSQY 88
Qy 98 -----LIEAIQKNATAFKGYAFKTYNYTL-----GADLTTPFGENQMVNSGKIFRYR 146
Db 89 LNNFOEKLNGSIPVNFSPENPLCFIKQWTPVIDAENADQLSSRGRLELFDLGRQLYOR 148
Qy 147 YKALARKIYVPIRAGSQRVATASAEKFIQFOSAKLADPGSQPHQASPVINIIPEG--S 204
Db 149 YKLFDSYVYDINTAEQERVESAKWFTYGLFGDMYKTN-----FLLISEGKAA 199
Qy 205 GYNNTLDHGTCTAFDSEL-----GDVEANFTALFAIPARLEADLPVTLTDEDVYV 259
Db 200 GANLSLMYNACPVFKDNFHNKATDAAHAVWRNIEFIVNRLAKYFSSYKLTINDVRS 259
Qy 260 LMDMCPFDVTARTSDATLSPPCALFTDHWITQYDYLQSLGKYGYGAGNPLGPAGQGVGF 319
Db 260 LFYICEYEIAIK-----DHSDFCSIFTPEFLNFEYDSDLDQAYGGPVSEWASTLGGAY 314
Qy 320 ANELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTMISIFFALGLYNGTK 379
Db 315 INNADSL-----RNYTNPDFDRK-----VFLAFTHDSNIIIPVEAALGFPPDIT 358
Qy 380 PLSTTSVESIEETDGYASWTVPFAARAYVEMMQCOAEPVRLVNDRVVPLHGC 439
Db 359 PQNPLTPDKNIYTSQKTSFVFPFAGNLTLEFFC-SDSKYVYRHLVNDRVVPLHGC 417
Qy 440 KLGR-----CKRDFVFEGLSFARSGGNWAECE 466
Db 418 PSGTSDGLCELQAYLNSPIRANSTNGISIF 448
RESULT 13
A25326
acid phosphatase (EC 3.1.3.2) precursor [similarity] - fission yeast (Schizosaccharom
C: Species: Schizosaccharomyces pombe
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C: Accession: A25326; T50405
R: Elliott, S.; Chang, C.; Schweingruber, M.E.; Schaller, J.; Rickli, E.E.; Carbon, J.
J. Biol. Chem. 261, 2936-2941, 1986
A: Title: Isolation and characterization of the structural gene for secreted acid phos
A: Reference number: A25326; MUID: 86140050
A: Accession: A25326
A: Molecule type: DNA
A: Residues: 1-453 <ELL>
A: Cross-references: GB:M11857; NID:g173422; PIDN:AAA35321.1; PID:g173423
R: Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 2000
A: Reference number: 225068
A: Accession: T50405
A: Molecule type: DNA
A: Residues: 1-453 <RIE>
A: Cross-references: EMBL:AL137099; PIDN:CA86857.1; GSPDB:GN00067; SPDB:SPBP4G3.02
A: Experimental source: strain 972h(-); clone p1 p4G3
C: Genes: phbB
A: Gene: phbB; SPDB:SPBP4G3.02
A: Map position: 2
C: Superfamily: yeast acid phosphatase
C: Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F: 68/Active site: Arg #status predicted
F: 69/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 12.1%; Score 297; DB 1; Length 453;
Best Local Similarity 24.8%; Pred. No. 4.3e-16;
Matches 99; Conservative 60; Mismatches 190; Indels 50; Gaps 9;

Qy 51 YSPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSA-----LIE 100
Db 43 HKPYF-----YGPSIDFFETCKIKQVHTLQRHGSRNPT--GGNAAFDAVGIANFOORLLNG 96
Qy 101 AIQKNATAFKGYAFKTYNYTL-----GADLTTPFGENQMVNSGKIFRYRKYKALARKIYVPF 157

Db 97 SVPIDYSGNELSFVPTWPIEAAADALSSGKVELFDMGRQFYERYHELNFNASTYN 156
QY 158 IRASGSDRVIAAEKFIIEGFSQAKLADPGSQPHQSPVINVIIPGSGYNNTLDHGTCTA 217
Db 157 IYTAQQRVDSALWYGYGMF-----GEDVHNFNTYILVSENATAGSNLSSSYNACPA 209
QY 218 FEDSELGDDVEANFTALFAPAIRLEADLPVTLTDEDVVVLMDCPDDTVARTSDATE 277
Db 210 SDADEFTTALAEARNVYPPRQRLNRYFNSVNTNDLNLGYCISYEIALQ-----D 264
QY 278 LSPFCALFTHDEWIOYDIQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTS 337
Db 265 YSEFCLENSVDFLNFYEGDLFSYGMGNSVKWGSIFGGAYANSLSL--RSVENNTQ 322
QY 338 TWHITLDSNPATFPLNATLYADFSHNTMISIFALGLYNGTKPLSTTSVESIEETDGYSA 397
Db 323 -----YQVFFAFTHDANIIPVETALGFFTDNTPENPLPTSYQVHSHMKA 366
QY 398 SWTPVFAARAYVEMMOCAKEPLRVLRVLRVPLHGC 436
Db 367 SEFVFPAGNLITELFQCEDSKI-YVRHLVNEEVFPLSDC 404

RESULT 14
S14119
acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
C:Accession: S14119; T40455
R:Yang, J.; Schweingruber, M.E.
Curr. Genet. 18, 269-272, 1990
A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schizosaccharomyces pombe
A:Reference number: S14119; MUID:91064763
A:Accession: S14119
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <YAN>
A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA40258.1; PID:g5007
R:Lyne, M.; Rajandream, M.A.; Barrall, B.G.; Baker, S.; Mungall, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21931
A:Accession: T40455
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-463 <LYN>
A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN00067; SPDB:SPBC428.03c
A:Experimental source: strain 972h-; cosmid c428
C:Genetics:
A:Gene: SPBC428.03c
A:Map position: 2
C:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 11.3%; Score 278.5; DB 2; Length 463;
Best Local Similarity 25.5%; Pred. No. 1.4e-14;
Matches 108; Conservative 65; Mismatches 187; Indels 63; Gaps 15;
QY 46 HLWGYG----PYFLADESAISPDVDCRVTFVQVLSRHGARYPT--SSASKAYSA-- 97
Db 35 HL-GTISVYHEPF-----NGPTTSFPESCAIKQVHLQLRHGSRNPTGDDTATDVSSAQ 88
QY 98 -----LIFAIQKNATAFKGYAFLKTYNTL---GADDITPFCENQMVNSGKIFKFR 146
Db 89 IDIFONKLLNGSYPNFSTPENPLFYVHWTFVIAKNAQDLSGRIELFDLGRQVFER 148
QY 147 YKALARKIVPFIIRASGRVIRASAEKFEIGFQSAKLADPGSQPHQSPVINVIIPGSGYNNTLDHGTCTA 217
Db 149 YVELFDTDYDINTAAQERVVDSAEWFSYGMFGDDMQNKTN-----FIVLPEDDSA 199
QY 205 GYNNTLDHGTCTAFDESELGDD--VEANFTA---LFAPAIRARLEADLPVTLTDEDVYV 259

Db 200 GANSLAMYSCPVYEDNNDIDENTEAAHTSWRNVLKPIANRLNKYFDSGYNLTVDVRS 259
QY 260 LMDMCPFDVTARTSDATELSPFCALFTHDEWIOYDIQSLGKYGYGAGNPLGPAQGVGF 319
Db 260 LYVICVYEIALRDN-----SDFCSLFTPSEFLNFYDSDLDYAYWGGPASEWASTLGGAY 314
QY 320 ANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHNTMISIFALGLYNGTK 379
Db 315 VNNLANNL-----RKGNNASDRK-----VFLAFTHDSQIIPVEAALGFEPDIT 358
QY 380 PLSTTSVESIEFTDGYSAISWTPVFAARAYVEMMOCAKEPLRVLRVLRVPLHGC 439
Db 359 PBHPLPTDKNIPTYSILKTSFVFPFAGNLITELFLC--SDNKYYVVRHLVNCQVVPVPLTDCGYG 417
QY 440 KLG 442
Db 418 PSG 420

RESULT 15
JE0369
histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JE0369
R:Mullaney, E.J.; Ullah, A.H.J.
Biochem. Biophys. Res. Commun. 251, 252-255, 1998
A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidopsis thaliana
A:Reference number: JE0369; MUID:99009256
A:Accession: JE0369
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <MUL>
C:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 6.1%; Score 149; DB 2; Length 465;
Best Local Similarity 22.3%; Pred. No. 0.00037;
Matches 101; Conservative 63; Mismatches 199; Indels 90; Gaps 22;
QY 35 DGGYQCFPEISHLMGTYSPYFSLAD--ESAI-SPDVDPDCRVTFVQVLSRHGARYPTSSA 91
Db 16 DGGF----DVRHHLSTVTRYSTSKDVTQNLIEGNSVPSECTPIHLNLVARHCTRSPTKKR 71
QY 92 SKAYSAL---TEAIQKNATAFK-----GKYAFLKTYNTLTCADDLTPFGENQMVN 138
Db 72 LREMESLAGRFKELVRDAEARKLPDKIPGWLQWK--SPWEGKVGKGLIRQGEDELYQ 129
QY 139 SGIKFYRRYKALARK---IVPFIIRASGSDRVIAAEKFIIEGFSQAKLADPGSQPHQASP 194
Db 130 LGIRVRERFPLFEDYHPDVYTIIRATQIPRASASAVAFGMGLFSEK-----GNL----- 179
QY 195 VINVIIPGSGYNNTLDHGTCTAFDESELGDDVEANFTALFAPAIRARLEADLPVTLTLD 254
Db 180 -----GPGNRNRA-----AVTSENKASDTKLRFEC--CQNTKSKYKAKEPAVDKLK 224
QY 255 EDVYVLMDC-----PDDTVARTSDATELSPPFCALFTHDEWIOYDIQSLGKYGYGAGNPL 311
Db 225 EPLVANKITASVAKRYDLKFTKQDITSSLMFLCKQVALLW-----TDDLEVLFLKGYGNSL 279
QY 312 GPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHNTMISIFFA 371
Db 280 NYKMGV-----PLLEDVLSHMBEAIKAREKLP--PGSYE-KARL--RFAHAETIIVPPFSC 330
QY 372 -LGILYNGTKPLSTTSVESIE-----ETDGYSAISWTPVFAARAYVEMMOCAKE 418
Db 331 LIGLFP-----LDGSEFEKIQKEKPLELPQPPKTRDRFGSTMAPFGGNNILVLYSCPAES 385
QY 419 BP--LVRLVLRVLRVPLHGC 442
Db 386 SPKYEQVQLHNEHPVAVPGC--DGKDFCPLDEF 416

Search completed: October 26, 2001, 15:10:29
Job time: 6335 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:18:03 ; Search time 29.88 Seconds
(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGFFVLLSTATLFGTSET.....DFVEGLSFARSGGNNAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	76.3	467	1	PHYA_ASPAW
2	1875	76.2	467	1	PHYA_ASPNG
3	1835	74.6	463	1	PHYB_EMENI
4	395	16.0	467	1	PPAB_YEAST
5	395	16.0	467	1	PPAC_YEAST
6	392	15.9	467	1	PPAS_YEAST
7	380	15.4	467	1	PPA3_YEAST
8	355	14.4	468	1	PPAL_PICPA
9	354	14.4	468	1	PPAD_YEAST
10	348.5	14.2	479	1	PHYB_ASPNG
11	347.5	14.1	479	1	PHYB_ASPAW
12	341.5	13.9	469	1	PPA5_KLULA
13	297	12.1	453	1	PPAL_SCHPO
14	278.5	11.3	463	1	PPA2_SCHPO
15	130	5.3	735	1	PPAX_CABEL
16	124.5	5.1	421	1	PPAL_MOUSE
17	114.5	4.7	423	1	PPAL_RAT
18	113	4.6	874	1	COPG_HUMAN
19	112	4.5	721	1	NICA_CABEL
20	102.5	4.2	423	1	PPAL_HUMAN
21	101.5	4.1	680	1	NCPR_CARMA
22	101	4.1	874	1	COPG_BOVIN
23	98.5	4.0	381	1	PPAP_RAT
24	96.5	3.9	275	1	UL11_HCMVA
25	96.5	3.9	386	1	PPAP_HUMAN
26	96	3.9	411	1	PPAY_CABEL
27	96	3.9	413	1	PPAW_CABEL
28	95.5	3.9	680	1	NCPR_CANTR
29	95	3.9	605	1	AMY_THECU
30	94.5	3.8	580	1	UL25_HSV11
31	94.5	3.8	602	1	PEHX_ERWCH
32	94.5	3.8	760	1	MCMT_SCHPO
33	94	3.8	754	1	MBPI_KLULA

ALIGNMENTS

RESULT 1

ID	PHYA_ASPAW	STANDARD;	PRT;	467 AA.
AC	P34753;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE			
DE	3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE			
DE	PHOSPHOHYDROLASE A).			
GN	PHYA OR PHY.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ALK0243;			
RX	MEDLINE=94040796; PubMed=8224894;			
RA	Piddington C.S., Houston C.S., Palcheimo M.T., Cantrell M.A.,			
RA	Miettinen-Oinonen A., Nevalainen H., Rambosk J.A.;			
RT	"The cloning and sequencing of the genes encoding phytase (phy) and			
RT	pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.			
awamori."				
RL	Gene 133:55-62(1993).			
CC	-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE			
CC	FROM PHYTASE.			
CC	-!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =			
CC	D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.			
CC	-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL: L02421; AAA16898.1; -			
DR	PIR: JN0889; JN0889.			
DR	HSSP: P34752; LIHP.			
DR	InterPro: IPR000560; -			
DR	Pfam: PF00328; acid_phosphat; 1.			
DR	PROSITE: PS00616; HIS.ACID.PHOSPHAT_1; 1.			
DR	PROSITE: PS00778; HIS.ACID.PHOSPHAT_2; 1.			
DR	Hydrolase; Glycoprotein; Signal.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	467	3-PHYTASE A.
FT	ACT_SITE	82	82	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE	361	361	PROTON DONOR (BY SIMILARITY).
FT	DISULFID	31	40	BY SIMILARITY.
FT	DISULFID	71	414	BY SIMILARITY.
FT	DISULFID	215	465	BY SIMILARITY.
FT	DISULFID	264	282	BY SIMILARITY.
FT	DISULFID	436	444	BY SIMILARITY.

34	93.5	3.8	2077	1	TEGU_HSV6G	p30002 human herpe
35	93.5	3.8	2077	1	TEGU_HSV6U	p52340 human herpe
36	92.5	3.8	417	1	AGP_PRORE	Q52309 providencia
37	92.5	3.8	844	1	YDZ5_SCHPO	013712 schizosacch
38	92	3.7	764	1	METE_SOLSC	Q42662 solenostemo
39	91	3.7	551	1	AMYB_THETU	p19584 thermoanaer
40	90	3.7	484	1	YODQ_YEAST	Q08193 saccharomyc
41	89.5	3.6	526	1	CLOS_CLOHI	p09870 clostridium
42	89.5	3.6	848	1	YAME_SCHPO	Q10061 schizosacch.
43	89.5	3.6	2208	1	POLN_MANCV	Q69014 manchester
44	89	3.6	823	1	AXL2_YEAST	p39288 saccharomyc
45	88.5	3.6	913	1	AGLU_BETVU	Q04931 beta vulgar


```

FT DISULFID 31 40
FT DISULFID 71 414
FT DISULFID 215 465
FT DISULFID 264 282
FT DISULFID 436 444
FT CARBOHYD 27 27
FT CARBOHYD 59 59
FT CARBOHYD 105 105
FT CARBOHYD 120 120
FT CARBOHYD 207 207
FT CARBOHYD 230 230
FT CARBOHYD 339 339
FT CARBOHYD 352 352
FT CARBOHYD 376 376
FT CARBOHYD 388 388
FT SEQUENCE 467 AA: 51086 MW: 88FE8F3584341D6D CRC64;

```

Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 1.3e-142;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

```

QY 1 MGVFVLLSIATLFGSTGTLGPRGNHSDTVDGGYQCFPEISHLMGTSPYFSLADE 60
Db 1 MGVSAVLLPLVLLSGVTSGLAVPASRNQSCDVTVDGGYQCFSETSHLWGYAPFESLANE 60
QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SWISEVPAGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
QY 121 YTLGADDTLTPGEGNOMVNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
Db 121 YSLGADDTLTPGEGNOMVNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
QY 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTCTAFEDSELGDDVEANFTALFAPAIR 240
QY 241 ARLEADLPVGLTDEDDVYVLLMDMCPDFTVARTSDATLSPFCALFTHDEWIQDYLOSIG 300
Db 241 QRLNDLSGVTLTDEDDVYVLLMDMCPDFTVARTSDATLSPFCALFTHDEWIQDYLOSIG 300
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPTATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPTATPLNATLYADES 360
QY 361 HDNMTISIFFALGYNKTKPLSTTSVESIETDGYASWTVPPAARAYVEMMOCQAEKEP 420
Db 361 HDNMTISIFFALGYNKTKPLSTTSVESIETDGYASWTVPPAARAYVEMMOCQAEKEP 420
QY 421 LVRVLYNDRVVPLHGCPCVADKLGCRKDDFVEGLSFARSGGNWAECEFA 467
Db 421 LVRVLYNDRVVPLHGCPCVADKLGCRKDDFVEGLSFARSGGNWAECEFA 467

```

RESULT 3
 PHYB_EMENI STANDARD; PRT; 463 AA.
 AC O00093;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 13-DEC-1998 (Rel. 37, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL-HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE B)
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98007872; PubMed=9349716;
 RA Pasamontes L., Halker M., Henriquez-Huecas M., Mitchell D.B.,

van Loon A.P.;
 "Cloning of the phytases from Emericella nidulans and the
 thermophilic fungus Talaromyces thermophilus";
 Biochim. Biophys. Acta 1353:217-223(1997).
 RL -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H₂O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

EMBL: U59803; AAB96871.1;
InterPro: IPR000560;
DR Pfam: PF00328; acid_phosphat_1;
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 463
FT ACT_SITE 80 80
FT ACT_SITE 81 81
FT ACT_SITE 357 357
FT CARBOHYD 26 26
FT CARBOHYD 41 41
FT CARBOHYD 103 103
FT CARBOHYD 118 118
FT CARBOHYD 203 203
FT CARBOHYD 226 226
FT CARBOHYD 331 331
FT CARBOHYD 335 335
FT CARBOHYD 372 372
FT SEQUENCE 463 AA; 51786 MW; ECC5827D1E1C82A2 CRC64;

```

Query Match 74.6%; Score 1835.5; DB 1; Length 463;
 Best Local Similarity 74.0%; Pred. No. 1.9e-139;
 Matches 345; Conservative 46; Mismatches 70; Indels 5; Gaps 3;

```

QY 1 MGVFVLLSIATLFGSTGTLGPRGNHSDTVDGGYQCFPEISHLMGTSPYFSLADE 60
Db 1 MGVFVLLSIATLFGSTGTLGPRGNHSDTVDGGYQCFPEISHLMGTSPYFSLADE 60
QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
QY 121 YTLGADDTLTPGEGNOMVNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
Db 121 YTLGADDTLTPGEGNOMVNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
QY 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTCTAFEDSELGDDVEANFTALFAPAIR 240
QY 241 ARLEADLPVGLTDEDDVYVLLMDMCPDFTVARTSDATLSPFCALFTHDEWIQDYLOSIG 300
Db 241 ARLEADLPVGLTDEDDVYVLLMDMCPDFTVARTSDATLSPFCALFTHDEWIQDYLOSIG 300
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPTATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPTATPLNATLYADES 360
QY 361 HDNMTISIFFALGYNKTKPLSTTSVESIETDGYASWTVPPAARAYVEMMOCQAEKEP 420
Db 361 HDNMTISIFFALGYNKTKPLSTTSVESIETDGYASWTVPPAARAYVEMMOCQAEKEP 420
QY 372 HDNMTISIFFALGYNKTKPLSTTSVESIETDGYASWTVPPAARAYVEMMOCQAEKEP 415

```

CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: U00029; AB69729.1; -
DR PIR: S48996; S48996.
DR HSP: P34752; 1HP.
DR SGD: S0001258; PHO12.
DR InterPro: IPR000560; -
DR Pfam: PF00328; acid_phosphat_1;
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17 BY SIMILARITY
FT CHAIN 18 467 ACID PHOSPHATASE PHO12.
FT ACT_SITE 75 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 337 337 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 467 AA; 52699 MW; BEC606CDF39B845B CRC64;

Query Match 16.0%; Score 395; DB 1; Length 467;
Best Local Similarity 26.0%; Pred. No. 3.7e-24;
Matches 119; Conservative 66; Mismatches 207; Indels 66; Gaps 13;

QY 24 PRGSHSCDVTGQYCFPELISHLWGYSPYFSLADSAISDPVDCRVTFVQLSRHG 83
DB 21 PLGLSDIDKIGTQTEIFPEFL---GGSGPYSPGYSRDLPSCEMKQVMVGRHG 76
QY 84 ARYPTSSASKAYSALIEAIOKNATAPKGYAFLK-----TYNYTL-----GADDL 128
DB 77 ERYPTVKAKSIMTWYKLSNYTQFSGALSFLNDYEFTRDTKNLEMETTLANSVNVL 136
QY 129 TPF-GENOMVNSGKIFRYRYKALARKIVFP-IRASGSDRVIAAEKFIQFQSAKLADPG 186
DB 137 NPYTGEMNAKRHARDFLAQYGVYVNETSFVATSTNSNRCHDTAQYFDG-----LGDKF 191
QY 187 SQPQASPVNVIPEGSYNTLDHCTCFADSELGDDVEANFTALFAPARLEAD 246
DB 192 NISLQ-----TISAESAGANTLSAHSACPWDD-DVNDILKKYDTKYLSGIAKRLNKE 245
QY 247 LPQVTLTDEDVYVLMDCPDTVARTSDATLSPFCALFTHDEWIDYQDYLOSLGKYGVG 306
DB 246 NKGLNLTSSDANTFFAWCAEINAR-----GYSDICNIFTKDELVRFSYQDLETYYQTG 300
QY 307 AGNPLGPAQGVGFANELIARLTSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMI 366
DB 301 PGYDVRSVGNANFNASVKLLKESEVDQ-----KWLSTHTDIDIL 342
QY 367 SIFPGLYNGTLPSTSTVESIEEDGYASWTVPFAARAYVEMQCAKEPLRVILV 426
DB 343 NYLTIGIIDDQNNLTAERHVPFMENT--FHRSWVVGARVYTEKFC--SNDTYVRYVI 398
QY 427 NDRVVLHGCAVDKLGCRKRDDE-----VEGLSFAR 457

DB 399 NDAVPLETCTGTGGSCSEINDEYGYAEKRVAGTDFLK 436
RESULT 6
PPA5_YEAST STANDARD; PRT; 467 AA.
ID AC P00635;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
GN PHO5 OR YBR093C OR YBR0814.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-26.
RX MEDLINE=83168913; PubMed=6300772;
RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
RT of repressible acid phosphatase contains a signal peptide.";
RL Nucleic Acids Res. 11:1657-1672(1983).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-45.
RX MEDLINE=85037940; PubMed=6093051;
RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
RT "Structural analysis of the two tandemly repeated acid phosphatase
RT genes in yeast.";
RL Nucleic Acids Res. 12:7721-7739(1984).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95208357; PubMed=7900426;
RA Manhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
RL Yeast 10:1363-1381(1994).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=87064474; PubMed=3537710;
RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
RA Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
RT cluster within the acid phosphatase multigene family of Saccharomyces
RL cerevisiae.";
RL Mol. Cell. Biol. 6:1855-1865(1986).
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01320; CAA24630.1; -
DR EMBL: X01079; CAA25555.1; -
DR EMBL: X78993; CAA55598.1; -
DR EMBL: Z35962; CAA85046.1; -
DR EMBL: X01080; CAA25556.1; -
DR PIR: S05795; PABYC.
DR PIR: B25241; B25241.
DR HSP: P34752; 1HP.
DR SGD: S0000297; PHO5.
DR InterPro: IPR000560; -
DR Pfam: PF00328; acid_phosphat; 1.
DR

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 103 103
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 456 456
 FT CARBOHYD 461 461
 FT CARBOHYD 461 461
 FT CONFLICT 36 36
 FT CONFLICT 130 130
 FT CONFLICT 294 294
 FT CONFLICT 446 446
 FT CONFLICT 462 462
 FT CONFLICT 466 466
 FT CONFLICT 466 466
 SQ SEQUENCE 467 AA; 52058 MW; DC3C9504BC2D3D0C CRC64;

Query Match 15.9%; Score 392; DB 1; Length 467;
 Best Local Similarity 25.6%; Pred. No. 6.5e-24;
 Matches 123; Conservative 64; Mismatches 216; Indels 78; Gaps 15;

QY 6 VLLSTATFGSTGALPGRNHSCOTVDGYOCFFPEISHLWCTSPYSLADESAISP 65
 DB 5 VVYSILASLANAGTI--PLGLADVDKIGTKDIFPL-----GGAPYISFPDGYISR 58
 QY 66 DVPDDCRVTFVQVLSRHGARYPTSASKAYSALIEATOKNATAFKGYAFLLK-TYNTTLG 124
 DB 59 DLPECEKQLQMGVGRHGERYPYTVSLAKTIKSTWYKLSNYTRQNGSLFLNDDYEFFIR 118
 QY 125 ADD-----LTPF-GENQVNSGKPYRYKALKARKIVFP-IPASGSDRVIA 168
 DB 119 DDDLEMETTFANSDVLPNPTGEMNAKRHRARDFLAQGYGVNENQTSFAVETSNSKRCHD 178
 QY 169 SAEKIEGFQSAKLADPGSQPHQASPVNLVII-----PEGSGYNNLTLDHGCTAFEDSEL 223
 DB 179 TQAYFIDGL-----GDQ-----FNITLQTVSEAESAGANTLSACNSCPAW-DYDA 222
 QY 224 GDDVEANFTALFAPAIRLEADLPGLVTLTDEDVYVLMDCPFDVTARTSDATSLSPFCA 283
 DB 223 NDDIIVNEYDTTVDLDAKRLNKENKGLNLTSTDASTLFSWCAFEV-----NAKGYSDVCD 277
 QY 284 LTHDEWTOYDLSGLKGYGAGNPLGPAAGVGFANELLARLTHSPVODHTSTNHTLD 343
 DB 278 IFTKDELHYSYQDLHTYHYHGGPYDIKISVGSNLFNASVKLLKQSEIQDQ----- 329
 QY 344 SNPATFPLNATLYADFSHNTMISITFFALGLYNGTKPLSTTSVESIETDGYASWTVPF 403
 DB 330 -----KVLWSFTHTDILNLTAGIIDDKNLTAEVVPMGNT--FHRSWYVPQ 377
 QY 404 ARAYVEMQCAEKEPLRVLVNDVPLHGCADVKGKCRKDDF-----VEGLSFA 456
 DB 378 GARVYTEKFCQ--SNDTYVRYVINDAVVPIETCTSGPGFSCEINDFYDAEKRVAAGTDFL 435
 QY 457 R 457
 DB 436 K 436

RESULT 7
 PPA3_YEAST
 ID PPA3_YEAST
 AC P24031;
 STANDARD; PRT; 467 AA.

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO3 OR YBR092C OR YBR0813.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=85037940; Pubmed=6093051;
 RX Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RA "Structural analysis of the two tandemly repeated acid phosphatase
 genes in yeast."; 12:7721-7739(1984).
 RT Nucleic Acids Res. 12:7721-7739(1984).
 RL [2]
 SEQUENCE FROM N.A.
 RP STRAIN=S288C;
 RC MEDLINE=95208357; Pubmed=7900426;
 RX Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
 RA "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X01080; CAA25557.1; -
 DR EMBL; X78993; CAA55597.1; -
 DR EMBL; Z35961; CAA85045.1; -
 DR PIR; S44674; PABYCC.
 DR HSP; P34752; 1IHP.
 DR SGD; S0000296; PHO3.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 103 103
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 456 456
 FT CARBOHYD 461 461
 FT CONFLICT 219 221
 SQ SEQUENCE 467 AA; 52776 MW; 05FEB80DEB41B0FF CRC64;

Query Match 15.4%; Score 380; DB 1; Length 467;

Best Local Similarity 26.3%; Pred. No. 5.9e-23;

Matches 114; Conservative 61; Mismatches 196; Indels 62; Gaps 11;

QY 49 GTYSPYSLADESAISPDDCRVTFVQVLSRHGARYPTSASKAYSALIEATOKNATA 108
 DB 42 GGAPYISFPDGYISRLPEGCEMQLARHGERTYPTYSKGTATMKTWYKLSNYTRQ 101

RESULT	8
PPAL_PICPA	
ID	PPAL_PICPA STANDARD; PRT; 468 AA.
AC	P52291;
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	ACID PHOSPHATASE PHOI PRECURSOR (EC 3.1.3.2).
GN	PHOI.
OS	Pichia pastoris (Yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Pichia.
OX	NCBI_TaxID=4922;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96001238; PubMed=7557473;
RA	Payne W.E., Gannon P.M., Kaiser C.A.;
RT	"An inducible acid phosphatase from the yeast pichia pastoris:
RT	characterization of the gene and its product.";
RL	Gene 163:19-26(1995)
CC	-I- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
CC	ALCOHOL + ORTHOPHOSPHATE.
CC	-I- INDUCTION: BY PHOSPHATE STARVATION.
CC	-I- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcements or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; U28658; AAB95503.1; ..
DR	HSSP; P34752; LIHP.
DR	InterPro; IPR000560; ..
DR	Pfam; PF00328; acid_phosphat; 1.
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW	Hydrolase; Glycoprotein; Signal; Multigene family.
FT	SIGNAL 1 22 POTENTIAL.
FT	CHAIN 23 468 ACID PHOSPHATASE PHOI.
FT	ACT_SITE 84 84 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).

RESULT 10
PHYB ASPNG

QY 217 AFDESELGDDVEANFTALFAPAIRARLEADLPGLTLDVYVLMDCPFDTVARTSDAT 276
 Db 218 TDNDQTTCDNLTYQLPQFVAA--ARLNSQNGMNLTSADVNLIVMASFELNAR----- 270
 QY 277 ELSPEALFTHDEWYQDYLSGLKYYGYGAGNPLGPAQGGVGFANELLARLTHSPVQDHT 336
 Db 271 PFSNINAFQTDEWYSGFYVEDLNYCYCAGPGCKKMAAVGAVYANASLTLLNQGP----- 325
 QY 337 SINHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLY--NGTKPLSTTSVESIEETDG 394
 Db 326 -----KEAGP-----LFFNFADHTNITPILAAAGVLIPNEDLPL-----DRVAFGNP 367
 QY 395 YSASWTVPFAARAYVEMMOCQA-----EKEPLRVLVNDRVVPVPHGC 436
 Db 368 YSIGNIVPMGGHLTIERLSCQATLSDAGTYVRLVINEAVLPFND 413

RESULT 11
 PHYB_ASPAW
 ID PHYB_ASPAW STANDARD; PRT; 479 AA.
 AC P34755;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (PH 2.5 OPTIMUM ACID PHOSPHATASE).
 GN PHYB OR APH.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=ALK0243;
 RX MEDLINE=94040796; PubMed=8224894;
 RA Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
 RA Miettinen-Oinonen A., Nevalainen H., Rambosk J.A.;
 RT "The cloning and sequencing of the genes encoding phytase (phy) and
 RT pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
 RT awamori".
 RL Gene 133:55-62(1993).
 RN 12
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=99264417; PubMed=10329192;
 RA Kostrewa D., Wyss M., D'Arcy A., van Loon A.P.;
 RT "Crystal structure of Aspergillus niger pH 2.5 acid phosphatase at
 RT 2.4-A resolution".
 RL J. Mol. Biol. 288:965-974(1999).
 CC 1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE.
 CC 1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC 1- SUBUNIT: HOMODIMER.
 CC 1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC
 CC EMBL; LO2420; AA16897.1;
 CC PIR; JN0890; JN0890.
 CC PDB; 1QFX; 19-APR-00.
 CC InterPro; IPR000560;
 CC Pfam; PF00328; acid_phosphat; 1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; 3d-structure.
 FT SIGNAL 1 19 BY SIMILARITY.

FT CHAIN 20 479 3-PHYTASE B.
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR.
 FT ACT_SITE 337 337 PROTON DONOR.
 FT DISULFID 71 387
 FT DISULFID 128 472
 FT DISULFID 216 441
 FT DISULFID 225 298
 FT DISULFID 413 421
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 479 AA; 52678 MW; 4F8E0F3778CCB08 CRC64;

Query Match 14.1%; Score 347.5; DB 1; Length 479;
 Best Local Similarity 29.1%; Pred. No. 2.5e-20;
 Matches 118; Conservative 54; Mismatches 171; Indels 63; Gaps 16;

QY 49 GTYSPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASK-AYSALIEATQKNAT 107
 Db 53 GYSESVSY---GIARDPPTSCVDQVIMVKRGERYPSPSAGKDIEEALAKVYSINTT 108
 QY 108 AFGKVAELKTYNYTL-----GADDLT-PF-CENQMVNSGKIFRYRYKAL--ARKIVPF 157
 Db 109 EYKGDIAFLNDWTYYVNECYNAETSGPYAGLLDAYNHGNDYKARYGHLNGETVVPF 168
 QY 158 IRASGSDRVIASAEKFIIEGFQSAKLADPGSQHQASPVINVII-PEGSYNNNTLDHGTC 216
 Db 169 F-SSGGRVETARKFEGEGFGYN-----YSTNAALNIISESEVMGADSLTP--TCD 217
 QY 217 AFDESELGDDVEANFTALFAPAIRARLEADLPGLTLDVYVLMDCPFDTVARTSDAT 276
 Db 218 TDNDQTTCDNLTYQLPQFVAA--ARLNSQNGMNLTSADVNLIVMASFELNAR----- 270
 QY 277 ELSPEALFTHDEWYQDYLSGLKYYGYGAGNPLGPAQGGVGFANELLARLTHSPVQDHT 336
 Db 271 PFSNINAFQTDEWYSGFYVEDLNYCYCAGPGCKKMAAVGAVYANASLTLLNQGP----- 327
 QY 337 SINHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLY--NGTKPLSTTSVESIEETDG 394
 Db 326 -----AGSLFFNFADHTNITPILAAAGVLIPNEDLPL-----DRVAFGNP 367

RESULT 12
 PPA5_KLULA
 ID PPA5_KLULA STANDARD; PRT; 469 AA.
 AC P52289;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHOS.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RA Feminin E.;
 RL Thesis (1995), University of Salamanca, Spain.
 CC 1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC 1- SUBCELLULAR LOCATION: SECRETED.
 CC 1- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).
 CC 1- PTM: GLYCOSYLATED DURING THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC 1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

STRAIN=97J2;
Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
-!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
-!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
ALCOHOL + ORTHOPHOSPHATE.
-!- SUBCELLULAR LOCATION: CELL WALL.
-!- INDUCTION: REPPRESSED BY THIAMINE.
-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch).

EMBL; X56939; CAA040258.1; - -
EMBL; AL034382; CAA22278.1; - -
PIR; S14119; S14119.
InterPro; IPR000560; - -
Pfam; PF00328; acid_phosphat; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
Hydrolase; Glycoprotein; Cell wall; Signal.

```

RESULT 15
PPAX_CAEEL
ID PPAX_CAEEL STANDARD; PRT; 755 AA.
AC Q09549; Q17843;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE ACID PHOSPHATASE F26C11.1 (EC 3.1.3.2).
EN F26C11.1.
GC
GN Caenorhabditis elegans.
OS
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RC Matthews P., Lloyd C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE

```

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z47072; CAA87370.1; -;
DR EMBL; Z54342; CAA87370.1; JOINED.
DR EMBL; Z54342; CAA91156.1; -;
DR EMBL; Z47072; CAA91156.1; JOINED.
DR WormPep; F26C11.1; CE05732.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein; Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
SQ SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.38; Score 130; DB 1; Length 755;
Best Local Similarity 21.9%; Pred. No. 0.012;
Matches 90; Conservative 55; Mismatches 130; Indels 136; Gaps 22;
QY 57 LADESAISPDVDDCRVTFVOVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFL 116
DB 332 LAIEKSFSEV-DNWKLEFVQTIWRHGD-----SALEGLFPISEKNWTFGG----- 378
QY 117 KTYNTLGGADLTFFGENQMYNSGKIFRYR-----KALARKIVPFIRASGSDRVI 167
DB 379 -----GLGELTPMGSEMNNLGTIFRRYVEDQQFLSHRYAAKET--YIRSTNLNRTI 429
QY 168 ASAEKFIEGFSAKIADPGSQPHQASPVINVIIEGSGYNNLTLD--HG----- 213
DB 430 ISAMSLLYG-----MFPPGAW-----NI---OGVDYPNDVDVMOQGGTFIPVHVDGID 473
QY 214 -----TCTAFED-----SELGDDVEANFTALFAIRARLEADLPVTLTDEDVYVL 260
DB 474 QCAVAQLCNCRRFOLEKWAEL-DEVKNATVAMI--ALNRRVAA---FYNVTDQP----- 523
QY 261 MDMCPFDTVARTSATELSPPC-----ALFTHDEWIDYDYLQSLGKYG-----YG 306
DB 524 -----EKENRYTDAWK-----CORNWFNDTMYQQLPWYNEDLYNEAQRTYAPFKRFTGN 573
QY 307 AGNP-----LGPAQGVGFANELIARLTHSPVQDHTSTNHTLD--SNPATPLN 352
DB 574 FGNRPSPVVDGIDIPQEVSTLQGGPLNLEIFEREKIRCVADAENCSDIVLKLKF--- 630
QY 353 ATLADFSHDNTMISIFFALGLYNGTKPLSTSVESIEETDGYGASWTVPF 403
DB 631 ---YAYSSHDQLVYALLVTLGI-----TDVVKTVDGWPDTSSSLTIEY 670

Search completed: October 26, 2001, 15:18:04
Job time: 450 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:04 ; Search time 57.24 Seconds
(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGVEVLLSLATLFGSTG.....DFEGLSFARSGNWAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*

```

1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2462	100.0	467	20 AAY43169	Consensus phytase-
2	2462	100.0	467	21 AAB20526	Consensus phytase
3	2462	100.0	467	21 AAY69568	Mutant phytase-1,
4	2443	99.2	467	21 AAB20532	Consensus phytase
5	2438	99.0	467	21 AAB20531	Consensus phytase
6	2414	98.1	467	20 AAW93382	Fungal consensus p
7	2408	97.8	467	20 AAW93380	Fungal phytase pro
8	2408	97.8	467	20 AAW93381	Fungal consensus p
9	2408	97.8	467	21 AAB20515	Consensus phytase
10	2408	97.8	467	21 AAY69558	Phytase-1, a conse
11	2407	97.8	467	20 AAW93383	Fungal consensus p

12	2405	97.7	467	20 AAW93384	Fungal consensus p
13	2399	97.4	467	20 AAW93385	Fungal consensus p
14	2398	97.4	467	20 AAY39906	Ascomycete Consens
15	2362	95.9	467	20 AAY43170	Consensus phytase-
16	2362	95.9	467	21 AAB20527	Consensus phytase
17	2362	95.9	467	21 AAY69569	Mutant phytase-10,
18	2360	95.9	467	21 AAB20534	Consensus phytase
19	2355	95.7	467	21 AAB20533	Consensus phytase
20	2342	95.1	467	21 AAB20524	Consensus phytase
21	2342	95.1	467	21 AAY69566	Phytase-10, a cons
22	2290	93.0	467	20 AAY43172	Consensus phytase-
23	2290	93.0	467	21 AAY69571	Phytase-7, a deriv
24	2287	92.9	467	21 AAB20529	Consensus phytase
25	2282	92.7	441	21 AAB20514	Consensus phytase
26	2216	90.0	441	21 AAB20523	Consensus phytase
27	2215	90.0	467	21 AAB20530	Consensus phytase
28	2215	90.0	467	21 AAY69572	Consensus phytase
29	2134	86.7	431	21 AAY69575	Consphyl2, a deriv
30	2134	86.7	431	21 AAB20535	Consensus phytase
31	2044	83.0	437	21 AAY69557	Initial consensus
32	2044	83.0	437	21 AAB20525	Consensus phytase
33	2032	82.5	467	21 AAY69567	Phytase-11, a cons
34	2012	81.7	467	20 AAY43171	Mutant Aspergillus
35	2012	81.7	467	21 AAB20528	A. fumigatus phyta
36	2012	81.7	467	21 AAY69570	Aspergillus fumiga
37	1956.5	79.5	424	21 AAB20536	Mutant Aspergillus
38	1956.5	79.5	424	21 AAY69565	Consensus phytase
39	1879	76.3	467	15 AAR46792	Initial consensus
40	1879	76.3	467	15 AAR46234	Phytase. Aspergil
41	1875	76.2	467	12 AAR11333	A. niger phytase.
42	1875	76.2	467	18 AAW15124	Recombinant chromo
43	1875	76.2	467	20 AAY39904	Aspergillus ficum
44	1867	75.8	465	19 AAW84356	Aspergillus fumiga
45	1867	75.8	465	20 AAY39905	A. fumigatus phyta

ALIGNMENTS

```

RESULT 1
AAY43169
ID AAY43169 standard; Protein; 467 AA.
AC
AC AAY43169;
XX
XX
DT 06-JAN-2000 (first entry)
XX
DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.
XX
KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
KW consensus sequence.
XX
OS Synthetic.
XX
XX W09948380-AL.
XX
XX 30-SEP-1999.
XX
XX 22-MAR-1999; 99WO-DK00154.
XX
XX 23-MAR-1998; 98DK-0000407.
XX
XX 19-JUN-1998; 98DK-0000806.
XX
XX 18-SEP-1998; 98DK-0001176.
XX
XX 22-JAN-1999; 99DK-0000091.
XX
XX 22-JAN-1999; 99DK-0000093.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Petersen S;
XX
XX WPI; 1999-591030/50.
XX
XX N-PSDB; AAZ31520.
XX

```

PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 9; 7lpp; English.

XX This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently. In addition to improved
 CC phytase-expressing transgenic plants, these plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 20; Length 467;
 Best Local Similarity 100.0%; Pred. No. 4e-239;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALPGRNHSCDVTVDGGYQCFFPEISHLMGTSPYFSLADE 60
 DB 1 mgffvllsiatlfgstgtalpgrnshscdvtvdggycfpeishlmgtspyfslade 60

QY 61 SAISDPVDDCRVTFVQLSRHGARYPTSSASKAYSALIEALQKNATAFKGYAFLLKTYN 120
 DB 61 saispdvddcrvtfvqlsrhgaryptssaskaysaliealqknatafkgyafllkty 120

QY 121 YTLGADDLTPFGENQMVNSGIKFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalar k ivpfir asgsdr v iasae k f i e g f q s a 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqphqaspvinviipgsgynnltldhgtctafedselgddveanftalfap 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDVTARTSDATLSPFCALFTHDEWIQDYQLQSLG 300
 DB 241 arleadlpvltltdedvvyvylm d c p f d v t a r t s d a t l s p f c a l f t h d e w i q d y q l q s l g 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarlthspvqdhstntnhtldsnpatfpplnatlyadfs 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMCOAKEP 420
 DB 361 hdn t m i s i f f a l g l y n g t k p l s t t s v e s i e e t d g y s a w t v p f a a r a y v e m m c o a k e p 420

QY 421 LVRVLVNDRVVPLHGCADVKGCRKRDDEFVGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndrvvplhgcadvkgcrkrddfvgl s f a r s g g n w a e c f a 467

RESULT 2

AAB20526

ID AAB20526 standard; Protein; 467 AA.

XX AAB20526;

AC AAB20526;

DT 05-DEC-2000 (first entry)

XX Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytase degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX W0200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI: 2000-491161/43.

XX N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 21; Length 467;
 Best Local Similarity 100.0%; Pred. No. 4e-239;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALPGRNHSCDVTVDGGYQCFFPEISHLMGTSPYFSLADE 60
 DB 1 mgffvllsiatlfgstgtalpgrnshscdvtvdggycfpeishlmgtspyfslade 60

QY 61 SAISDPVDDCRVTFVQLSRHGARYPTSSASKAYSALIEALQKNATAFKGYAFLLKTYN 120
 DB 61 saispdvddcrvtfvqlsrhgaryptssaskaysaliealqknatafkgyafllkty 120

QY 121 YTLGADDLTPFGENQMVNSGIKFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalar k ivpfir asgsdr v iasae k f i e g f q s a 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqphqaspvinviipgsgynnltldhgtctafedselgddveanftalfap 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDVTARTSDATLSPFCALFTHDEWIQDYQLQSLG 300
 DB 241 arleadlpvltltdedvvyvylm d c p f d v t a r t s d a t l s p f c a l f t h d e w i q d y q l q s l g 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarlthspvqdhstntnhtldsnpatfpplnatlyadfs 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMCOAKEP 420
 DB 361 hdn t m i s i f f a l g l y n g t k p l s t t s v e s i e e t d g y s a w t v p f a a r a y v e m m c o a k e p 420

QY 421 LVRVLVNDRVVPLHGCADVKGCRKRDDEFVGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndrvvplhgcadvkgcrkrddfvgl s f a r s g g n w a e c f a 467

RESULT 3
 AAY69568
 ID AAY69568 standard; protein; 467 AA.
 XX
 AC AAY69568;
 XX
 DT 19-APR-2000 (first entry)
 XX
 DE Mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.
 XX
 KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutein.
 XX
 OS Aspergillus terreus 9A1.
 OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger T213.
 OS Aspergillus fumigatus ATCC13073.
 OS Aspergillus fumigatus ATCC13072.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Emericella nidulans.
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Synthetic.
 XX
 EH Key Location/Qualifiers
 FT Peptide 1..26
 FT /note= "Phytase signal peptide from Aspergillus terreus
 FT cbs16.46"
 FT Protein 27..467
 FT /note= "Mature phytase-1-thermo[8]-Q50T-K91A"
 XX
 FN EP969089-AL.
 XX
 PD 05-JAN-2000.
 XX
 XX 23-JUN-1999; 99EP-0111949.
 XX
 PR 29-JUN-1998; 98EP-0111960.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Brugger R, Lehmann M, Wyss M;
 DR WPI: 2000-099429/09.
 DR N-PSDB; AAZ59715.
 XX
 PT New stabilized enzyme formulation, useful for feed compositions for
 PT monogastric animals -
 XX
 PS Example 5; Fig 19; 101pp; English.
 XX
 CC The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is

CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-1
 CC consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a
 CC temperature optimum and melting point 7 degrees Celsius higher than that
 CC of phytase-1 (AAY69538).
 XX
 XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 21; Length 467;
 Best Local Similarity 100.0%; Pred. No. 4e-239;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSLATLFGSTGTALGPRGNHSCDFTVDGOCFPEISHLWCTYSPYSLADE 60
 DB 1 mgvfvlslatlfgstgtalgrgnshscdftvdgocfpeishlwtyspyfslade 60
 QY 61 SATSPDVPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saisdpvdpdcrrvtfvqlsrhgaryptssaskaysalieleaigknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIKASGSDRVIASAEKFIKGFOSA 180
 DB 121 ytlgaddltpfgenqmvnsgkifryrykalkarkivpfirkasgsdrviasaekfiegfosa 180
 QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAPEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqphqaspvinviipgsgynntldhgtctapedselgddveanfalfapair 240
 QY 241 ARLEADLPVGTUTDEVDVYLMDCPFDVTARTSDATELSPFCALFTHDEWIDYDYLQSLG 300
 DB 241 arleadlpvgtutdevdvyldmcpfdvtartsdatelspfcalfthdewidydyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFPFGALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 DB 361 hdntmisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcakep 420
 QY 421 LVRVLNDRVPLHGCANDKLGRCRDRDDFVEGLSFARSGGNWAECEFA 467
 DB 421 lvrvlndrvplhgcavdklgrckrdddfveglsfarsggnwaecefa 467

RESULT 4
 AAB20532
 ID AAB20532 standard; Protein; 467 AA.
 XX
 AC AAB20532;
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.
 XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; reaction rate;
 KW substrate binding; position specificity; phytate cleavage pattern;
 KW food; feed; phytate; manure.
 XX
 OS Synthetic.
 XX
 PN WO2000/43503-A1.
 XX
 PD 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.
 PF

XX 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX
 XX (NOVO) NOVO NORDISK AS.
 XX
 PI Lehmann M;

XX WPI; 2000-491161/43.
 DR N-PSDB; AAA73291.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 23a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 99.2%; Score 2443; DB 21; Length 467;
 Best Local Similarity 99.4%; Pred. No. 3.3e-237;
 Matches 464; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGPRGNSHSCDTVDGGYQCFPEISHLWCTYSPYFSLADE 60
 DB 1 mgvfvlslsiatlfgstgaltgprgnshscdtvdggycfpeishlwnctytpyfslade 60
 QY 61 SAISPDPVDCRVTFFQVLSRHGARYPTSSKAYSALIEAIQKNATAFKGYAFKLTYN 120
 DB 61 saispdpvdkcrvtffqvlsrhgaryptssaskaysalieaiqknatafkgyafkltyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 121 ytlgaddltpfgnqmvnsgkifryrykalkarkivpfirasgsdrviasaekfiqfqs 180
 QY 181 KLADPGSOPHOASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsophoaspvinviipgsgynntldhgtctafedstlgddveanftalfapair 240
 QY 241 ARLEADLPGLVTLDEDDVYVLMDCPFDVTARTSDATSELPFCALFTHDEMIVQDYQLQSLG 300
 DB 241 arleadlpvltldeddvylmdcpfdvtartsdatselpfcalfthdewiqdyqlqslg 300
 QY 301 KYCYGAGNPLGPAQGVGFANELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarlthspvqdhtstnhtldsnpatplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNTKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCQAEKEP 420
 DB 361 hdnTmisiffalglyntkplsttsvesietdgyaswtvpfaarayvemmocqaekep 420
 QY 421 LVRVLVNDVRVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndvrvpplhgcavdkgcrkddfveglsfarsggnwaecefa 467

RESULT
 AAB20531

ID AAB20531 standard; Protein; 467 AA.

XX AAB20531;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO2000043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 99.0%; Score 2438; DB 21; Length 467;
 Best Local Similarity 99.1%; Pred. No. 1e-236;
 Matches 463; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGPRGNSHSCDTVDGGYQCFPEISHLWCTYSPYFSLADE 60
 DB 1 mgvfvlslsiatlfgstgaltgprgnshscdtvdggycfpeishlwnctytpyfslade 60
 QY 61 SAISPDPVDCRVTFFQVLSRHGARYPTSSKAYSALIEAIQKNATAFKGYAFKLTYN 120
 DB 61 saispdpvdkcrvtffqvlsrhgaryptssaskaysalieaiqknatafkgyafkltyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 121 ytlgaddltpfgnqmvnsgkifryrykalkarkivpfirasgsdrviasaekfiqfqs 180
 QY 181 KLADPGSOPHOASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsophoaspvinviipgsgynntldhgtctafedstlgddveanftalfapair 240

Db 181 kladpgsqhgapvlnvipegsgyntldhglctafedstlsgddveanftalfapair 240
 QY 241 ARLEADLPGLTDRDVLVMDMCPEDVARTSDATLSPFCALFTHDEWTDYDYLQSLG 300
 Db 241 arleadlpgvltltdedvvyldmcpfdkvarstsdatslpcalfthdewtdydlqslg 300
 QY 301 KYGGAGNPLGPAGVGGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyggagagnplgpagvggfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIETDGYSSASWTVPFAARAYVEMMOCQAEKEP 420
 Db 361 hdnmtisiffalglngtkplsttsvesietdgyssaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLNDRVPLHGCVAVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 Db 421 lvrvlndrvvplhgcavdklgrckrddfvglsgfarsgggnwaecefa 467

RESULT 6
 AAW93382
 ID AAW93382 standard; Protein; 467 AA.
 AC AAW93382;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal consensus phytase protein mutant Q50T.
 DE
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX
 PN EP897985-A2.
 XX
 PD 24-FEB-1999.
 XX
 PF 15-JUL-1998; 98EP-0113176.
 XX
 PR 24-JUL-1997; 97EP-0112688.
 XX
 PA (HOFF) HOFFMANN LA ROCHE AG F.
 XX
 PI Lehmann M;
 XX
 DR WPI; 1999-134647/12.
 XX
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 XX
 PS Claim 9; Page -; 30pp; English.
 XX
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.
 XX
 SQ Sequence 467 AA;

Query Match 98.1%; Score 2414; DB 20; Length 467;
 Best Local Similarity 98.1%; Fred. No. 2.7e-234;
 Matches 458; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MGWVVLISATLFGSTGALPRGNSHSCDVTGQYQCFPEISHLWGTSPYFSLADE 60

Db 1 mgwfvllsatiifgstsaltgprgnshscdvtggyqcfpeishlwtgyspyfslade 60
 QY 61 SATSPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFILKTYN 120
 Db 61 saispdvddcrvtfvqvlshrgharyptsskskaysalieleaigknatafkgyafilktyn 120
 QY 121 YTLGADDLPPFENQMVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFIQFOSA 180
 Db 121 ytlgaddlppfengqmvnsgikfryrykalarlkiwpffirasgsdrviasaekfiegfosa 180
 QY 181 KLADPGSQHQAQSPVNLVPIEGSGYNNNTLDHGTCTATPDESELGDDVDVANFTALFAPAIR 240
 Db 181 kladpgsqhqaqspvnlvpipegsgynntldhgtctatpdeselgddveanftalfapair 240
 QY 241 ARLEADLPGLTDRDVLVMDMCPEDVARTSDATLSPFCALFTHDEWTDYDYLQSLG 300
 Db 241 arleadlpgvltltdedvvyldmcpfdkvarstsdatslpcalfthdewtdydlqslg 300
 QY 301 KYGGAGNPLGPAGVGGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyggagagnplgpagvggfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIETDGYSSASWTVPFAARAYVEMMOCQAEKEP 420
 Db 361 hdnmtisiffalglngtkplsttsvesietdgyssaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLNDRVPLHGCVAVDKLGCRKRDDEVEGLSFARSGGNWAECEFA 467
 Db 421 lvrvlndrvvplhgcavdklgrckrddfvglsgfarsgggnwaecefa 467

RESULT 7
 AAW93380
 ID AAW93380 standard; Protein; 467 AA.
 AC AAW93380;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal phytase protein consensus DNA.
 DE
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX
 PN EP897985-A2.
 XX
 PD 24-FEB-1999.
 XX
 PF 15-JUL-1998; 98EP-0113176.
 XX
 PR 24-JUL-1997; 97EP-0112688.
 XX
 PA (HOFF) HOFFMANN LA ROCHE AG F.
 XX
 PI Lehmann M;
 XX
 DR WPI; 1999-134647/12.
 XX
 DR N-PSDB; AAX23022.
 XX
 PT Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 XX
 PS Claim 8; Fig 2; 30pp; English.
 XX
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses
 CC phytase to valuable feed additives, with a fully defined amino acid
 CC sequence given in the specification, or variant or muten. The method is

CC useful for improving protein properties by altering their sequence. The
 CC consensus protein and mutin are useful in food, feed or pharmaceutical
 CC compositions. This sequence represents the consensus phytase protein used
 CC in the method of the invention.

XX
 XX Sequence 467 AA;

Query Match 97.8%; Score 2408; DB 20; Length 467;
 Best Local Similarity 97.9%; Pred. No. 1.1e-233;
 Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTDDGGYQCPEISHLWGTYSYFSLADE 60
 DB 1 mgvfvlslatlfgstgtalgrgnshscdvtddggycpeshlwlwgtysyfslede 60
 QY 61 SAISPDPVDDCRVTFFVOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdpddcrvtffvqlsrhgaryptsskskaysalialieaiqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgengmvmnsgikfyrrykalarkivpfirasgsdrviasaekfiegtfqa 180
 QY 181 KIADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kiadpgsqphqaspvldvlipegsynntldhgtctafedseelgddveanftalfapair 240
 QY 241 ARLEADLPVGLTDEDVYVLMDCPFTVARTSDATELSPFCAETHDEWIOYDYQLQSLG 300
 DB 241 arleadlpvgtltddedvvyldmcpftvartsdatselspfcalthdewioydyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpqagvgfaneliarltrspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMQCAKEP 420
 DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyaswtpfagarayvemmqcqaekp 420
 QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDDEFGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndrvvplhgcavdklgrckrddfveglfsarsggnwaecfa 467

RESULT 8
 AAW93381
 ID AAW93381 standard; Protein; 467 AA.

XX
 AC AAW93381;

DT 11-JUN-1999 (first entry)

XX Fungal consensus phytase protein mutant Q50L.

XX phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 XX feed additive; variant; mutin; feed; pharmaceutical.

OS Fungi.
 OS Synthetic.

XX EP897985-A2.

XX 24-FEB-1999.

XX 15-JUL-1998; 98EP-0113176.

XX 24-JUL-1997; 97EP-0112688.

XX (HOFF) HOFFMANN LA ROCHE AG F.

XX Lehmann M;

XX

DR WPI; 1999-134647/12.

XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences

XX Claim 9; Page -: 30pp; English.

XX This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or mutin. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and mutin are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.

XX Sequence 467 AA;

Query Match 97.8%; Score 2408; DB 20; Length 467;
 Best Local Similarity 97.9%; Pred. No. 1.1e-233;
 Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTDDGGYQCPEISHLWGTYSYFSLADE 60
 DB 1 mgvfvlslatlfgstgtalgrgnshscdvtddggycpeshlwlwgtysyfslede 60
 QY 61 SAISPDPVDDCRVTFFVOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdpddcrvtffvqlsrhgaryptsskskaysalialieaiqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgengmvmnsgikfyrrykalarkivpfirasgsdrviasaekfiegtfqa 180
 QY 181 KIADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kiadpgsqphqaspvldvlipegsynntldhgtctafedseelgddveanftalfapair 240
 QY 241 ARLEADLPVGLTDEDVYVLMDCPFTVARTSDATELSPFCAETHDEWIOYDYQLQSLG 300
 DB 241 arleadlpvgtltddedvvyldmcpftvartsdatselspfcalthdewioydyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpqagvgfaneliarltrspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMQCAKEP 420
 DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyaswtpfagarayvemmqcqaekp 420
 QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKRDDEFGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndrvvplhgcavdklgrckrddfveglfsarsggnwaecfa 467

RESULT 9
 AAB20515
 ID AAB20515 standard; Protein; 467 AA.

XX
 AC AAB20515;

XX 05-DEC-2000 (first entry)

XX Consensus phytase SEQ ID NO:16.

XX phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 XX temperature stability; pH profile; temperature profile; reaction rate;
 XX specific activity; substrate specificity; substrate cleavage pattern;
 XX substrate binding; position specificity; phytate degradation rate;
 XX food; feed; phytate; manure.

001 nansmsrlllaiglyngtaprsttsvesieetdgyssawtvpfgarayvemmqaekep 420

feed in the form of phytate phosphorus, monogastric animals are

CC to utilise this form of phosphate, resulting in the addition of extra
CC phosphate to the feed of such animals. Phytase enhances the nutritional
CC value of plant material without the need for adding additional phosphate
CC to the feed. The level of phosphate pollution in the environment is
CC reduced by adding phytase to animal feed, as the animal can make use of
CC the inorganic phosphate liberated from phytate phosphorus using the
CC enzyme. The phytase formulation of the invention has an improved
CC thermostability and can therefore remain stable during long-term storage
CC and can withstand feed processing methods such as extrusion, expansion,
CC and pelleting. The present sequence represents a consensus sequence,
CC designated phytase-1, which was derived from the mature phytase
CC sequences from a variety of fungi (AA69544-V69556), and additionally
CC contains the Aspergillus terreus csb116.46 signal peptide at the
CC N-terminus.

XX Sequence 467 AA;
SQ

Query Match 97.8%; Score 2408; DB 21; Length 467;
Best Local Similarity 97.9%; Pred. No. 1.1e-233;
Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTDDGGYQCFPEISHLWGTYSYFSLADE 60
DB 1 mgvfvlslsiatlfgstgtalgrgnshscdvtddggycfpeishlwgysyfslede 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
DB 61 saispdvddcrvtfvqlsrhgarpytssaskaysalieaiqknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQGFQSA 180
DB 121 ytlgaddltpfgenqmvnsgkifryrykalkarkivpfirasgsdrviasaekfieqfqs 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 kladpgsqphqaspvdiviipsgsynntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEADLPVTLTDEDEVVYLMDCPFDTVARTSDATELSPFCALFTHDEWIQDYLSLG 300
DB 241 arleadlpvltldedevvylmdcpfdtvartsdatelspfcalfthdewiqdylyslg 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCAKEP 420
DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyasaswtvpfagarayvemmqc 420

QY 421 LVRVLVNDVRVPLHGCVDKLGCRKRDDEFVGLSFARSGGNWAECEFA 467
DB 421 lvrvlvndrvvplhgcavdklgcrkrddefvglsfarsggnwaecefa 467

RESULT 11
AAW93383
ID AAW93383 standard; Protein; 467 AA.
XX
XX AAW93383;
AC
XX
XX 11-JUN-1999 (first entry)
DT
XX Fungal consensus phytase protein mutant Q50G.
DE
DE Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
KW feed additive; variant; mutein; feed; pharmaceutical.
KW
XX Fungl.
OS
OS Synthetic.
XX
XX EP897985-A2.
PN
XX

PD 24-FEB-1999.
XX
XX PF 15-JUL-1998; 98EP-0113176.
XX
XX PR 24-JUL-1997; 97EP-0112688.
XX
XX PA (HOFF) HOFFMANN LA ROCHE AG F.
XX
XX PI Lehmann M;
XX
XX DR WPI; 1999-134647/12.
XX
XX PT preparation of a consensus protein, especially a phytase - using
XX programs to compare evolutionary similarity of sequences
XX
XX PS Claim 9; Page -; 30pp; English.
XX

CC This invention describes a novel process for the preparation of a
CC consensus protein. The specific example given in the specification is
CC that of a fungal phytase (myo-inositol hexakisphosphate) which
CC hydrolyses phytase to valuable feed additives, with a fully defined amino
CC acid sequence given in the specification, or variant or mutein. The
CC method is useful for improving protein properties by altering their
CC sequence. The consensus protein and mutein are useful in food, feed or
CC pharmaceutical compositions. This sequence is a mutant phytase protein of
CC the invention which does not appear in the specification but has been
CC created from the protein represented in AAW93380.

XX Sequence 467 AA;
SQ

Query Match 97.8%; Score 2407; DB 20; Length 467;
Best Local Similarity 97.9%; Pred. No. 1.4e-233;
Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTDDGGYQCFPEISHLWGTYSYFSLADE 60
DB 1 mgvfvlslsiatlfgstgtalgrgnshscdvtddggycfpeishlwgysyfslede 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
DB 61 saispdvddcrvtfvqlsrhgarpytssaskaysalieaiqknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQGFQSA 180
DB 121 ytlgaddltpfgenqmvnsgkifryrykalkarkivpfirasgsdrviasaekfieqfqs 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 kladpgsqphqaspvdiviipsgsynntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEADLPVTLTDEDEVVYLMDCPFDTVARTSDATELSPFCALFTHDEWIQDYLSLG 300
DB 241 arleadlpvltldedevvylmdcpfdtvartsdatelspfcalfthdewiqdylyslg 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCAKEP 420
DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyasaswtvpfagarayvemmqc 420

QY 421 LVRVLVNDVRVPLHGCVDKLGCRKRDDEFVGLSFARSGGNWAECEFA 467
DB 421 lvrvlvndrvvplhgcavdklgcrkrddefvglsfarsggnwaecefa 467

RESULT 12
AAW93384
ID AAW93384 standard; Protein; 467 AA.
XX
XX AAW93384;
AC

XX 11-JUN-1999 (first entry)
 XX Fungal consensus phytase protein mutant Q50T-Y51N.
 DE
 XX
 XX Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 KW
 XX Fungi.
 OS Synthetic.
 OS
 PN EP897985-A2.
 XX 24-FEB-1999.
 PD 15-JUL-1998; 98EP-0113176.
 XX 24-JUL-1997; 97EP-0112688.
 XX (HOFF) HOFFMANN LA ROCHE AG F.
 PR Lehmann M;
 PI
 DR WPI; 1999-134647/12.
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 XX
 PS Claim 9; Page -: 30pp; English.
 PS
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.
 CC
 XX Sequence 467 AA;
 SQ

 Query Match 97.7%; Score 2405; DB 20; Length 467;
 Best Local Similarity 97.9%; Pred. No. 2.2e-233;
 Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCFPEISHLWGTYSPIYFLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshtcdtdvggyqcfpeishlwtgtnspysfslde 60
 QY 61 SAISPDPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNATAFKGYAFILKTYN 120
 Db 61 saispdpddcrvtfvqlsrhgaryptssaskaysalieleatqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKIEGFQSA 180
 Db 121 ytlgaddltpfgnqmvnsgikfryrykalkarkivpfirasgsdrviasaekiefqfqa 180
 QY 181 KIADPGSQPHQASPVNLVITPEGSYNNLTDLGCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 kladpgsqphqaspvnlvltpegsyynnltldgctafedselgddveanftalfapair 240
 QY 241 ARLEADLPVTLTDEDVYLMDCPFDTVARTSDATELSPFCALTFHDEWQYDYQLSLG 300
 Db 241 arleadlpvltltdedvylmdcpcfdtvartsdatelspfcaltfhdewqdyqlslg 300
 QY 301 KYGAGNPLGPGAGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyygagnplgpgagvgfaneliarltrspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGKPLSTTSVESIEETDGYSGASWTVPFAARAYVEMMQCAKEP 420

Db 361 hdnsmisiffalglngtaplstsvesieetdgyaswtvpfagarayvemmqcqaekp 420
 QY 421 LVRVLVNDRVVPLHGCAYDKLGRCKRDDRVEGLSFARSGNNAECFA 467
 Db 421 lvrvlvndrvvplhgcaydklgrckrddrddvfglsfarsggnwaecfa 467
 RESULT 13
 AAW93385
 ID AAW93385 standard; Protein; 467 AA.
 XX
 AC AAW93385;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal consensus phytase protein mutant Q50L-Y51N.
 XX
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX EP897985-A2.
 PN 24-FEB-1999.
 PD 15-JUL-1998; 98EP-0113176.
 XX 24-JUL-1997; 97EP-0112688.
 PR (HOFF) HOFFMANN LA ROCHE AG F.
 PA Lehmann M;
 PI
 DR WPI; 1999-134647/12.
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 XX
 PS Claim 9; Page -: 30pp; English.
 PS
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.
 CC
 XX Sequence 467 AA;
 SQ

 Query Match 97.4%; Score 2399; DB 20; Length 467;
 Best Local Similarity 97.6%; Pred. No. 8.8e-233;
 Matches 456; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCFPEISHLWGTYSPIYFLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshtcdtdvggyqcfpeishlwtgtnspysfslde 60
 QY 61 SAISPDPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNATAFKGYAFILKTYN 120
 Db 61 saispdpddcrvtfvqlsrhgaryptssaskaysalieleatqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKIEGFQSA 180
 Db 121 ytlgaddltpfgnqmvnsgikfryrykalkarkivpfirasgsdrviasaekiefqfqa 180

181 KLDGSGPHQASPVINVIIPGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240
 181 KLDGSGPHQASPVINVIIPGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240
 241 ARLEADLPVTLTDEDVYVLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 241 ARLEADLPVTLTDEDVYVLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPNATLYADFS 360
 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPNATLYADFS 360
 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467
 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467

RESULT 14

AA39906
 ID AA39906 standard; Protein; 467 AA.

XX AC AA39906;

XX 07-DEC-1999 (first entry)

XX Ascomycete consensus phytase protein sequence.

XX Phytase; variant; enzyme; phosphorus liberation; phytase substrate;

XX Phytate level reduction; animal manure; food preparation;

XX soy processing; inositol manufacture.

XX Synthetic.

XX WO9949022-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-DK00153.

XX 23-MAR-1998; 98DK-0000407.

XX 19-JUN-1998; 98DK-0000806.

XX 18-SEP-1998; 98DK-0001176.

XX 22-JAN-1999; 99DK-0000091.

XX (NOVO) NOVO-NORDISK AS.

XX Svendsen A;

XX WPI; 1999-580444/49.

XX N-PSDB; AA27423.

XX New variant phytase enzymes, used for liberating phosphorus from a
 phytase substrate, for reducing phytate levels in animal manure and in
 feed and food preparations

XX Claim 26; Fig 9a-d; 14lpp; English.

XX This sequence represents the ascomycetes consensus phytase sequence.

XX The invention relates to variant phytase enzymes with specific amino acid
 substitutions for improved properties. The phytase variants can be used
 for liberating phosphorus from a phytase substrate. They can be used for
 reducing phytate levels in animal manure. They can be used in feed or
 food preparations. The phytase DNA can also be used to produce transgenic
 plants which can be used in feeds or foods. The phytase variants can also
 be used in soy processing and in the manufacture of inositol or
 derivatives. The phytase variants can have altered activities such as pH
 stability, temperature stability, pH profile, temperature profile,
 specific activity (in particular in relation to pH and temperature),
 substrate specificity, substrate cleavage pattern, substrate binding,

CC position specificity, the velocity and level of release of phosphate
 from corn, reaction rate, phytase degradation rate and end level of
 released phosphate reached.
 XX SQ Sequence 467 AA;

Query Match 97.4%; Score 2398; DB 20; Length 467;

Best Local Similarity 97.4%; Pred No. 1,le-232;

Matches 455; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGALGPRGNHSCDTVDGGYOCFPEISHLMCTGTPYPSLSADE 60

Db 1 mgfvvllsiatlfgstgaglgprgnshscdtvdggyqcfpeishlmwgyqpylslede 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKTYN 120

Db 61 saispdvddcrvtfvqvlsrhgaryptsskaysalieleaigknatafkyafiktyN 120

QY 121 YTLGADDLTPFGENOMVNSGIRFYRYKALARKIYVPIFRASGSDRVIASAEKFIQFQSA 180

Db 121 ytlgaaddltpfgenomvnsigikfyrkyalarakiyvfirasgsdrviasaekfiqfqsA 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240

Db 181 kladpgsqphqaspvivdiipgsgynntldhgctafedselgddveanftalfapair 240

QY 241 ARLEADLPVTLTDEDVYVLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300

Db 241 arleadlpvgvltddedvvyldmcmptetvartsdatselcpalfthdewrgydlqslg 300

QY 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPNATLYADFS 360

Db 301 kyygaggnplgpaqgvgfaneliarltrspvqdhstnhtltdsnpatfpnatlyadfs 360

QY 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420

Db 361 hdnsmisiffalglyngtkplstsvesieetdgyaswtvpfagarayvemmqcqaekp 420

QY 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467

Db 421 lvrvlvndvrvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

RESULT 15

AA43170

ID AA43170 standard; Protein; 467 AA.

XX AC AA43170;

XX 06-JAN-2000 (first entry)

XX Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.

XX Phytase; animal feed preparation; thermostable phytase; transgenic plant;

XX consensus sequence.

XX Synthetic.

XX WO9948380-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-DK00154.

XX 23-MAR-1998; 98DK-0000407.

XX 19-JUN-1998; 98DK-0000806.

XX 18-SEP-1998; 98DK-0001176.

XX 22-JAN-1999; 99DK-0000091.

XX 22-JAN-1999; 99DK-0000093.

XX (NOVO) NOVO-NORDISK AS.

XX PA

XX XX

PI Petersen S;
XX
DR WPI; 1999-591030/50.
DR N-PSDB; AAZ31521.
XX
PT Preparing animal feed using a thermostable phytase
XX
PS Example 3; Fig 10; 71pp; English.
XX
CC This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
CC The invention relates to a process for preparing animal feed by
CC agglomerating feed ingredients with a thermostable phytase, which is
CC added before or during agglomeration. The thermostable phytase is useful
CC for expression in transgenic plants. These plants are useful in the
CC preparation of animal feed itself. The thermostable phytase allows animal
CC feed to be produced more efficiently, in addition to improved
CC phytase-expressing transgenic plants. These plants provide a feed
CC ingredient and a feed additive (phytase) simultaneously.
XX
SQ Sequence 467 AA;

Query Match 95.9%; Score 2362; DB 20; Length 467;
Best Local Similarity 95.5%; Pred. No. 4.7e-229;
Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFWLLSIATLFGSTGTALGPRGNHSCDVTGQYQCFPEISHLWGTYSYFSLADE 60
Db 1 mgfwllsiatlfgstgtalgrgnshcdvtvggyqcfpeishlwgtyssffslade 60

QY 61 SAISPDVPDCRVTFFQVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFLKTYN 120
Db 61 saispdvpcrvtffqvlshrhgaryptssaskaysalieaioknatarkgyafllkty 120

QY 121 YTLGADLTLPFGENQWNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltlpgfnqwnsgkifryrykalkarkivpfiragsgsdrviasaekfieqf 180

QY 181 KLADPGSOPHOASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 kladpgsophsopvpinviiipgsgynnltdhgtctafedselgddveanftavfapp 240

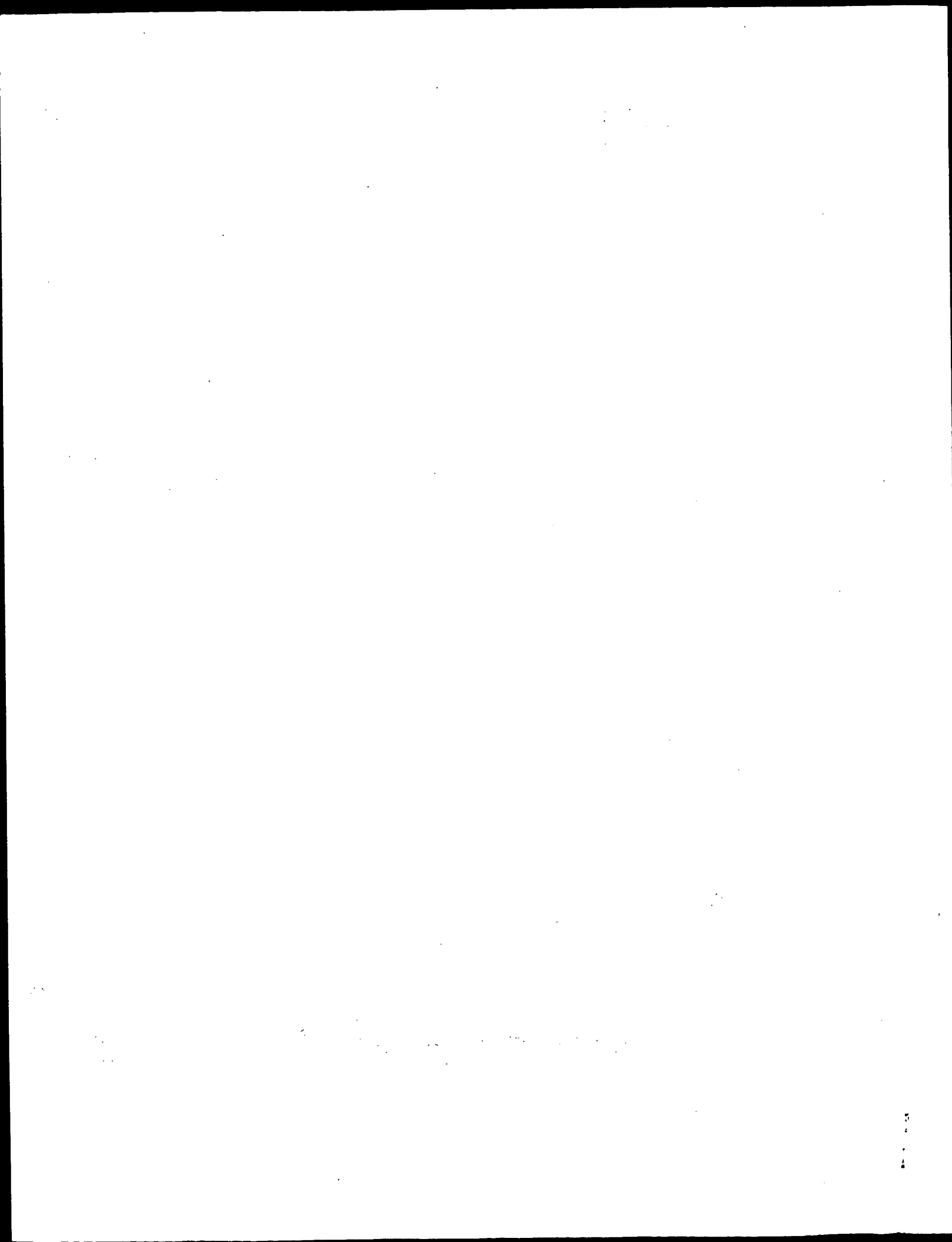
QY 241 ARLEADLPVTLTDEDDVYVLMDCPFDVTARTSDATSELSPECALFTHDEWIOYDYLSLG 300
Db 241 arleahlpvtnltdedvnlmdcpfdvtartsdatselqspfcldfthdewigdylyq 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
Db 301 kygygagnplgpaqgvgfvneliarltshspvqdhstnhtldsnpatfplnatlyad 360

QY 361 HDNMTISIFPALGLXNKTPLTSTVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
Db 361 hdnmtisifalglxngtkplstsviesieetdgyaswtvpfaarayvemmqqceake 420

QY 421 LVRVLNDRVPLHGCADVCKLGRKRDDEFGLSFARSGGNNAECFA 467
Db 421 lvrvlndrvplhgcadvcklgrckrddfveglsfarsggnweecfa 467

Search completed: October 26, 2001, 16:40:04
Job time: 4955 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:46 ; Search time 32.57 seconds
(without alignments)
295.232 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGFFVLLSLATLFGTSGT.....DFVEGLSFARSGNNAECFA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6C.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282	92.7	441	4	US-09-121-425-1
2	2270	92.2	467	4	US-09-121-425-2
3	1879	76.3	467	1	US-07-923-724-8
4	1879	76.3	467	2	US-08-609-426A-8
5	1879	76.3	467	2	US-08-374-652C-2
6	1875	76.2	467	1	US-08-151-574-32
7	1875	76.2	467	1	US-08-146-424-20
8	1875	76.2	467	1	US-08-693-709-2
9	1875	76.2	467	2	US-08-419-448-32
10	1875	76.2	467	4	US-08-819-825-3
11	1875	76.2	467	4	US-09-163-642-3
12	1860	75.5	467	4	US-09-155-855-3
13	1818	73.8	443	4	US-09-155-855-1
14	1813	73.6	443	4	US-09-155-855-2
15	1334	54.2	475	4	US-08-819-825-2
16	1334	54.2	475	4	US-09-163-642-2
17	802.5	32.6	443	3	US-08-993-359-30
18	798.5	32.4	439	3	US-08-993-359-24
19	798.5	32.4	439	3	US-09-221-654-2
20	798.5	32.4	439	3	US-08-989-358A-2
21	795.5	32.3	453	3	US-08-993-359-22
22	781	31.7	442	3	US-08-993-359-28
23	778	31.6	442	3	US-08-993-359-26
24	363	14.7	468	1	US-07-627-539G-2
25	358.5	14.6	446	1	US-07-627-539G-7
26	347.5	14.1	479	1	US-07-923-724-2
27	347.5	14.1	479	2	US-08-609-426A-2

28	347.5	14.1	479	2	US-08-374-652C-4	Sequence 4, Appli
29	170.5	6.9	449	3	US-08-680-506-7	Sequence 7, Appli
30	148.5	6.0	92	3	US-08-993-359-32	Sequence 32, Appli
31	122	5.0	113	1	US-08-241-853-8	Sequence 8, Appli
32	122	5.0	113	2	US-08-850-917-8	Sequence 8, Appli
33	119	4.8	24	2	US-08-374-652C-32	Sequence 32, Appli
34	118	4.8	318	3	US-08-680-506-3	Sequence 3, Appli
35	111	4.5	113	1	US-08-241-853-10	Sequence 10, Appli
36	111	4.5	113	2	US-08-850-917-10	Sequence 10, Appli
37	96.5	3.9	386	1	US-08-758-213-1	Sequence 1, Appli
38	96.5	3.9	386	2	US-08-692-787-48	Sequence 48, Appli
39	96.5	3.9	386	4	US-09-097-199-48	Sequence 48, Appli
40	96.5	3.9	515	2	US-09-146-283-2	Sequence 2, Appli
41	96.5	3.9	515	3	US-08-579-823A-2	Sequence 2, Appli
42	96.5	3.9	515	4	US-09-344-195-2	Sequence 2, Appli
43	95.5	3.9	1732	2	US-08-570-311-10	Sequence 10, Appli
44	95.5	3.9	1732	2	US-08-353-485-10	Sequence 10, Appli
45	93.5	3.8	20	1	US-07-923-724-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-1

Query Match	92.7%	Score	2282	DB	4	Length	441
Best Local Similarity	97.7%	Pred. No.	1.1e-230				
Matches	431	Conservative	3	Mismatches	7	Indels	0
						Gaps	0
QY	27	NSHSCDVTDDGGYQCFPELSHLMGTSPYFSLADESAISPDVDDCRVTFVQVLSRHGARY	86				
Db	1	NSHSCDVTDDGGYQCFPELSHLMGTSPYFSLADESAISPDVDDCRVTFVQVLSRHGARY	60				
QY	87	PTSSASKAYSALIEAIQKNATAFKGYAFKLTNYNTLGADDLTPFGENOMVNSGKFKYRR	146				
Db	61	PTSSASKAYSALIEAIQKNATAFKGYAFKLTNYNTLGADDLTPFGENOMVNSGKFKYRR	120				
QY	147	YKALARKIVPFRASGSDRVITASAEKFTGQSADKADPGSOPHOASPVINVIIEGSGY	206				
Db	121	YKALARKIVPFRASGSDRVITASAEKFTGQSADKADPGSOPHOASPVINVIIEGSGY	180				
QY	207	NNTLDHGTCFAFDESELGDDVEANFTALFAFARLEADLPVTLTDEDVYVLMDCPF	266				
Db	181	NNTLDHGTCFAFDESELGDDVEANFTALFAFARLEADLPVTLTDEDVYVLMDCPF	240				
QY	267	DTVARTSDATSLSPFCALFTHDEWIQDYDLSLQKYYGYGAGNPLGPAQGVGFANELIAR	326				
Db	241	ETVARTSDATSLSPFCALFTHDEWIQDYDLSLQKYYGYGAGNPLGPAQGVGFANELIAR	300				
QY	327	LTHSPVDHTSTNHTLDSNPATFLNATLYADFSDHNTMISIFFALGLYNGTKPLSTTSV	386				

```
Db 301 LTRSPQDHTSTNHTLDSNPATFPLNATLYADFSDHNSMISIFALGLYNGTAPLSTTSV 360
Qy 387 ESIEETDGYASWTVPFAARAYVEMMOCAEKEPLVRVLVNDVRVPLHGCADVCKLGRCKR 446
Db 361 ESIEETDGYASWTVPFAARAYVEMMOCAEKEPLVRVLVNDVRVPLHGCADVCKLGRCKR 420
Qy 447 DDFVEGLSFARSGGNWAECEFA 467
Db 421 DDFVEGLSFARSGGNWAECEFA 441

RESULT 2
US-09-121-425-2
; Sequence 2, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2

Query Match 92.2%; Score 2270; DB 4; Length 467;
Best Local Similarity: 89.7%; Pred. No. 2.1e-229;
Matches 437; Conservative 3; Mismatches 7; Indels 40; Gaps 2;

Qy 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGTVPYFSLADE 60
Db 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGTVPYFSLADE 60
Qy 61 SAISPDVDCRVTFVQVLSRHGARYPTSSAKYASALIEAIQKNATAPKGYAFKVTN 120
Db 61 SAISPDVDCRVTFVQVLSRHGARYPTSSAKYASALIEAIQKNATAPKGYAFKVTN 120
Qy 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
Db 101 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 160
Qy 181 KLADPGSQPHQASPVIN-----VIIPGSGYNNTLDHGTCTAFED 220
Db 161 KLADPGSQPHQASPVIDLIEAIQKNATAPKGYAFKVIIPGSGYNNTLDHGTCTAFED 220
Qy 221 SELGDDVEANFTALFAIRLEADLPVTLTDEVDVYLMDCPDTVARTSDATFELSP 280
Db 221 SELGDDVEANFTALFAIRLEADLPVTLTDEVDVYLMDCPDTVARTSDATFELSP 280
Qy 281 FCALFTHDEWIQDYLSGLKYYGAGNPLGPAQGVGFANELIARLTHTSPVODHTSTNH 340
Db 281 FCALFTHDEWRQDYLSGLKYYGAGNPLGPAQGVGFANELIARLTHTSPVODHTSTNH 340
Qy 341 TLDSNPATFPLNATLYADFSDHNTMISIFALGLYNGTAPLSTTSVESIEETDGYASWT 400
Db 341 TLDSNPATFPLNATLYADFSDHNSMISIFALGLYNGTAPLSTTSVESIEETDGYASWT 400
Qy 401 VPFAARAYVEMMOCAEKEPLVRVLVNDVRVPLHGCADVCKLGRCKRDFVEGLSFARSGG 460
Db 401 VPFAARAYVEMMOCAEKEPLVRVLVNDVRVPLHGCADVCKLGRCKRDFVEGLSFARSGG 460
Qy 461 NWAECEFA 467
Db 461 NWAECEFA 467
```

```
Db 461 NWAECEFA 467

RESULT 3
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8

Query Match 76.3%; Score 1879; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 2.1e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGTVPYFSLADE 60
Db 1 MGVSAPVLLPLLLAGLAVPASRNOSTCDTVDQGTQCFSETSHLWQVAPFSLANE 60
Qy 61 SAISPDVDCRVTFVQVLSRHGARYPTSSAKYASALIEAIQKNATAPKGYAFKVTN 120
Db 61 SAISPDVDCRVTFVQVLSRHGARYPTSSAKYASALIEAIQKNATAPKGYAFKVTN 120
```


QY 121 YTLGADLLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGEGFQSA 180
Db 121 YSLGADLLTPFGELVNSGKIFRYEYSLTRNIIPIRSGSSRVIASGEKFTGEGFQST 180
QY 181 KLADPGSQHOASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLKDPRAQPGQSSPKIDVIVISEASSNNTLDPGCTVEFDESELADTVEANFTAFAPAIR 240
QY 241 ARLEADLPGLVTLTDEDVYLMDCMCPDFTVARTSDATELSPFCALFTHDEWIQYDYLQSLG 300
Db 241 ORLENDLSGVTLTDEVTYLMDCMSFDTISTVDTKLSPCDLFTHDEWIHYDYLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGHGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HDNTWISIFFALGYNKTKPLSTTSVESIEBETDGYASWTVPFAARAYVEMMQCAKEP 420
Db 361 HDNGIISILFALGYNKTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAKEP 420
QY 421 LVRVLNDRVPLHGCAYDKLGRCKRDDFVEGLSFARSGGNWAECSA 467
Db 421 LVRVLNDRVPLHGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 4

US-08-609-426A-8
; Sequence 8, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.

; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-426A-8

Query Match 76.3%; Score 1879; DB 2; Length 467;

Best Local Similarity 75.6%; Pred. No. 2.1e-188;

Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTSGTALGPRGNHSHCDTVDGGYQCFPEISHLMGTSPYFSLADE 60
Db 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSTCDTVDGQYQCFSETSHLWGOYAPFSLANE 60
QY 61 SAISDPDPPDCRVTFQVLSRHGARYPTSSASKAYSALIEIAIQKNATAFKGYAFKTYN 120
Db 61 SAISDPDPAGCRVTFQVLSRHGARYPTESKGYKSALIEETQQNVTTFDGKYAFKTYN 120
QY 121 YTLGADLLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGEGFQSA 180
Db 121 YSLGADLLTPFGELVNSGKIFRYEYSLTRNIIPIRSGSSRVIASGEKFTGEGFQST 180
QY 181 KLADPGSQHOASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLKDPRAQPGQSSPKIDVIVISEASSNNTLDPGCTVEFDESELADTVEANFTAFAPAIR 240
QY 241 ARLEADLPGLVTLTDEDVYLMDCMCPDFTVARTSDATELSPFCALFTHDEWIQYDYLQSLG 300
Db 241 ORLENDLSGVTLTDEVTYLMDCMSFDTISTVDTKLSPCDLFTHDEWIHYDYLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGHGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HDNTWISIFFALGYNKTKPLSTTSVESIEBETDGYASWTVPFAARAYVEMMQCAKEP 420
Db 361 HDNGIISILFALGYNKTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAKEP 420
QY 421 LVRVLNDRVPLHGCAYDKLGRCKRDDFVEGLSFARSGGNWAECSA 467
Db 421 LVRVLNDRVPLHGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 5

US-08-374-652C-2
; Sequence 2, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON

```

; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-374-652C-2

```

```

Query Match 76.3%; Score 1879; DB 2; Length 467;
Best Local Similarity 75.6%; Pred. No. 2.le-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVEVLLSLATLFGSTGTALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPIYFSLADE 60
DB 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSCDVTVDGGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAFKGYAFLKTYN 120
DB 61 SAISDPVAGCRVTFQAQLSRHARGARYPTESKGYKYSALIEEQNVTTFDGKYAFLKTYN 120

QY 121 YTLGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFQYQYSLRNIIPIFIRSSGSRVIAASGEKFIQFQST 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 KLKDPRAQPGQSSPKIDVVIASESSNNLTDPGCTVFEDESLADTVEANFTATFAPSIR 240

QY 241 ARLEADLPGVTLTDEDVYVLMDCPFDVARTSDATELSPPFCALFTHDEWIOYDYLQSLG 300
DB 241 ORLENDLSGVTLTDEVTYVLMDCSFDTISTSTVDTKLSPPFCDLFTHDEWIHYDYLQSLK 300

QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNAILYADF 360
DB 301 KYYGAGNPLGPTQGVGVANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADF 360

QY 361 HONTMITSIFALGLYNGTKPLSTTSVESTTEETDGSASWTVPFAARAVEMQCAQEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMQCAQEP 420

QY 421 LVRVLVNDRVPLHGCVDKLCRKRDDFEGLSFARSGNWAECFA 467
DB 421 LVRVLVNDRVPLHGCPIDALGCRTRDSEFVRGLSFARSGGDWAECFA 467

```

```

RESULT 6
US-08-151-574-32
; Sequence 32, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selden
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-151-574-32

```

```

Query Match 76.2%; Score 1875; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVEVLLSLATLFGSTGTALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPIYFSLADE 60
DB 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSCDVTVDGGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAFKGYAFLKTYN 120
DB 61 SVISPEVAGCRVTFQAQLSRHARGARYPTDSKGYKYSALIEEQNVTTFDGKYAFLKTYN 120

QY 121 YTLGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFQYQYSLRNIIPIFIRSSGSRVIAASGEKFIQFQST 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 KLKDPRAQPGQSSPKIDVVIASESSNNLTDPGCTVFEDESLADTVEANFTATFAPSIR 240

QY 241 ARLEADLPGVTLTDEDVYVLMDCPFDVARTSDATELSPPFCALFTHDEWIOYDYLQSLG 300
DB 241 ORLENDLSGVTLTDEVTYVLMDCSFDTISTSTVDTKLSPPFCDLFTHDEWIHYDYLQSLK 300

```

```

QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNNHTLDSNPATFPPLNATLYADFS 360
DB 301 KYYGAGNPLGPTQGGVYANELIARLTHSPVHDDTSSNHTLDSNPATFPPLNATLYADFS 360
QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIETDGYASASVTVPFAARAYVEMMQCAKEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSAWTVPFASRLYVEMMQCAKEP 420
QY 421 LVRVLNDRVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
DB 421 LVRVLNDRVPLHGCPVDALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 7
US-08-146-424-20
; Sequence 20, Application US/08146424
; Patent No. 5593963
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,424
; FILING DATE: 02-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KENNEDY, BILL
; REGISTRATION NUMBER: 33,407
; REFERENCE/DOCKET NUMBER: 44615-20011.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-146-424-20

Query Match 76.2%; Score 1875; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-186;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVSFVLLSIATLFGSTGALCPNGSHSCDTPDGGYQCFPEISHLWGTSPYFSLADE 60
DB 1 MGVSFVLLPLLYLGLSVTGLAVPASRNSQCDTPDGGYQCFSETSHLWGTAPFFSLANE 60
QY 61 SAISFDVDDCRVTFVQLSRHGRARYPTSSASKAYSALIEAIQKNATAFKGKAFKTYN 120
DB 61 SVISPEVAGCRVTFVQLSRHGRARYPTSSASKAYSALIEAIQKNATAFKGKAFKTYN 120
QY 121 YTLGADLLTPFGENQWNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180

```

```

DB 121 YSLGADLLTPFGEQELVNSGKIFRYQYSELTRNIVPFFIRSSSSRVIASGKFIQFQST 180
QY 181 KLADPGSQHQSPVNTVIPIPGSYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 KLKDPRAQPGQSPKLDVVISEASSNNLTDPGTCTVFEDSELADTVEANFTATVPISR 240
QY 241 ARLEADLPQVTLTDEDDVYVLMDCPFDIVARTSDATELSFPFCALTHDEWTDYQLQSLG 300
DB 241 QRLNDLSGVTLTDEVTYVLMDCSPFDITSTVDTKLSPPCDLTHDEWINYDYLQSLK 300
QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNNHTLDSNPATFPPLNATLYADFS 360
DB 301 KYYGAGNPLGPTQGGVYANELIARLTHSPVHDDTSSNHTLDSNPATFPPLNATLYADFS 360
QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIETDGYASASVTVPFAARAYVEMMQCAKEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSAWTVPFASRLYVEMMQCAKEP 420
QY 421 LVRVLNDRVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
DB 421 LVRVLNDRVPLHGCPVDALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 8
US-08-693-709-2
; Sequence 2, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

```

FEATURE:
 NAME/KEY: Signal Sequence
 LOCATION: 1...23
 OTHER INFORMATION:
 US-08-693-709-2

Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 5.6e-188;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTDDGGYQCPEISHLWCTYSPYFSLADE 60
 DB 1 MGVSAVLLPLLYLGSVTSGLAVPASRNQSCDVTDDGGYQCPEISHLWCTYSPYFSLADE 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSGALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 SVISPEVPACRVTFQAQLSRHGARYPTDSKGGKYSALIEEIQONATTFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFFRASGSDRVIASAEKFEIGFQSA 180
 DB 121 YSLGADDLTPFGELVNSGKIFRYRYKALARKIVPFFRASGSDRVIASAEKFEIGFQST 180

QY 181 KLADPGSOPHOASPVNTIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 KLKDPRAQCGQSPKIDVVISASSNNTLDPCTCTVEDSELADTVEANFTATFVPSIR 240

QY 241 ARLEADLPGLVTLDEDDVYLMDCPFDVARTSDATELSPCALFTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVTLTDEVTYLMDCSFDITSTVDTKLSPCDFLTHDEWINDYQLQSLK 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
 DB 301 KYGHGAGNPLGTPQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMISIFLALGNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMQCAEKEP 420
 DB 361 HDNGIISILFALGNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCVDKLGRCRDRDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLVNDRVVPLHGCVPDVALGRCRDRDSFVRGLSFARSGGDWAECEFA 467

RESULT 9

US-08-419-448-32
 Sequence 32, Application US/08419448
 Patent No. 5863533

GENERAL INFORMATION:

APPLICANT: Robert F.M. Van Gorcom
 APPLICANT: Willem Van Hartingsveldt
 APPLICANT: Petrus A. Van Paridon
 APPLICANT: Annemarie E. Veenstra
 APPLICANT: Rudolf G.M. Luttin
 APPLICANT: Gerardus Seiten
 TITLE OF INVENTION: Cloning and Expression of Microbial
 TITLE OF INVENTION: Phytase
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 SOFTWARE:
 CURRENT APPLICATION NUMBER: US/08/419,448

FILING DATE: 10-APR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 24615-20026.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-887-1500
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-419-448-32

Query Match 76.2%; Score 1875; DB 2; Length 467;

Best Local Similarity 75.6%; Pred. No. 5.6e-188;

Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTDDGGYQCPEISHLWCTYSPYFSLADE 60
 DB 1 MGVSAVLLPLLYLGSVTSGLAVPASRNQSCDVTDDGGYQCPEISHLWCTYSPYFSLADE 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSGALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 SVISPEVPACRVTFQAQLSRHGARYPTDSKGGKYSALIEEIQONATTFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFFRASGSDRVIASAEKFEIGFQSA 180
 DB 121 YSLGADDLTPFGELVNSGKIFRYRYKALARKIVPFFRASGSDRVIASAEKFEIGFQST 180

QY 181 KLADPGSOPHOASPVNTIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 KLKDPRAQCGQSPKIDVVISASSNNTLDPCTCTVEDSELADTVEANFTATFVPSIR 240

QY 241 ARLEADLPGLVTLDEDDVYLMDCPFDVARTSDATELSPCALFTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVTLTDEVTYLMDCSFDITSTVDTKLSPCDFLTHDEWINDYQLQSLK 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
 DB 301 KYGHGAGNPLGTPQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMISIFLALGNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMQCAEKEP 420
 DB 361 HDNGIISILFALGNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCVDKLGRCRDRDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLVNDRVVPLHGCVPDVALGRCRDRDSFVRGLSFARSGGDWAECEFA 467

RESULT 10

US-08-819-825-3

Sequence 3, Application US/08819825

Patent No. 5866118

GENERAL INFORMATION:

APPLICANT: Berkka, Randy M.
 APPLICANT: Ray, Michael W.
 APPLICANT: Klotz, Alan V.
 TITLE OF INVENTION: Polypeptides Having Phytase Activity
 TITLE OF INVENTION: And Nucleic Acids Encoding Same
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 58661180 No. 58661180disk of No. 5866118th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-819-825-3

```

```

Query Match          76.2%; Score 1875; DB 2; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

```

```

QY 1 MGVEFVLLSIATLFGSTGALPGRGNHSCDVTVDGGYQCFPEISHLWGTSPYFSLADE 60
DB 1 MGVSALLPLLYLLSGVTGLAVPASRNSQCDTVDGQYCFSETSHLWGOYAPFSLANE 60
QY 61 SATSPDVPDCRVTFVQVLSRHGARYPTSSAKYASALIEAIOKNATAFKGYAFLKTYN 120
DB 61 SVISPEVPAGCRVTFQAQVLSRHGARYPTDSKGKYSALIEEIQONATTEDGKYAFLKTYN 120
QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPPIRAGSDRVIASAEKFTIEGQSA 180
DB 121 YSLGADDLTPFGEOLVNSGKIFQRYESLTRNIVPPIRSGSSRVIASGKFTIEGFQST 180
QY 181 KLADPGSQPHQASPVINVIPEGGSYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 KLADPRAQPGQSSPKIDVISEASSNNTLDPGTCTVFEDSELADTVEANFTATFVPSIR 240
QY 241 ARLEADLPVTLTDEDVYLMDCMCPDTVARTSATLSPFCALFTHDEWIQYDYLQSLG 300
DB 241 QRLNDLSGVTLTDEVTYLMDCMCSPTISTSTVDTKLSPFCOLFTHDEWIYDYLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYGYGAGNPLGTPQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASATVTPFAARAYVEMMOCAKEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTIVENITQDGFSSAWTVPFASRLYVEMMOCAQEP 420
QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
DB 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSEFVRGLSFARSGGDWAECEFA 467

```

```

RESULT 11
US-09-163-642-3
; Sequence 3, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSER: No. 6221644o No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-163-642-3

```

```

Query Match          76.2%; Score 1875; DB 4; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

```

```

QY 1 MGVEFVLLSIATLFGSTGALPGRGNHSCDVTVDGGYQCFPEISHLWGTSPYFSLADE 60
DB 1 MGVSALLPLLYLLSGVTGLAVPASRNSQCDTVDGQYCFSETSHLWGOYAPFSLANE 60
QY 61 SATSPDVPDCRVTFVQVLSRHGARYPTSSAKYASALIEAIOKNATAFKGYAFLKTYN 120
DB 61 SVISPEVPAGCRVTFQAQVLSRHGARYPTDSKGKYSALIEEIQONATTEDGKYAFLKTYN 120
QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPPIRAGSDRVIASAEKFTIEGQSA 180
DB 121 YSLGADDLTPFGEOLVNSGKIFQRYESLTRNIVPPIRSGSSRVIASGKFTIEGFQST 180
QY 181 KLADPGSQPHQASPVINVIPEGGSYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 KLADPRAQPGQSSPKIDVISEASSNNTLDPGTCTVFEDSELADTVEANFTATFVPSIR 240
QY 241 ARLEADLPVTLTDEDVYLMDCMCPDTVARTSATLSPFCALFTHDEWIQYDYLQSLG 300
DB 241 QRLNDLSGVTLTDEVTYLMDCMCSPTISTSTVDTKLSPFCOLFTHDEWIYDYLQSLK 300
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYGYGAGNPLGTPQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASATVTPFAARAYVEMMOCAKEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTIVENITQDGFSSAWTVPFASRLYVEMMOCAQEP 420
QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
DB 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSEFVRGLSFARSGGDWAECEFA 467

```

```

RESULT 12
US-09-155-855-3

```

; Sequence 3, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; SOFTWARE: PatentIn Ver. 2.0
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-3

Query Match 75.5%; Score 1860; DB 4; Length 467;
Best Local Similarity 74.18; Pred. No. 2.1e-186;
Matches 346; Conservative 48; Mismatches 73; Indels 0; Gaps 0;
QY 1 MGVFVLLSIATLFGSTGTCALCPRGNSHSDTVDGGYOCPEISHLWGTSPYFSLADE 60
DB 1 MGVSAVLLPLXLLSGVTSGLAVPASRNGSTCDTVDQGYOCFSETSHLWGVAPFESLANK 60
QY 61 SAISDPDPCDRTVQVLSRHGARYPTSSKAYSALIEAIQKNATFAKGYAFLKTYN 120
DB 61 SAISDPDPCDRTVQVLSRHGARYPTSSKAYSALIEAIQKNATFAKGYAFLKTYN 120
QY 121 YTLGADDTLPPFENOMVNSGKIFRYKALARKIVPFTIRASGSDRVIASAEKFTGQSA 180
DB 121 YSLGADDTLPPFENOMVNSGKIFRYKALARKIVPFTIRASGSDRVIASAEKFTGQSA 180
QY 181 KLADPGSQPHQASPVINIIPEGSGYNNLTLDHGCTAFEDSELGDDVDFANFTALFAPAIR 240
DB 181 KLADPGSQPHQASPVINIIPEGSGYNNLTLDHGCTAFEDSELGDDVDFANFTALFAPAIR 240
QY 241 ARLEADLPVTLTDEVDVYLMDCFPDVARSDATELSPCALTFHDEWIOYDYLQSLG 300
DB 241 ARLEADLPVTLTDEVDVYLMDCFPDVARSDATELSPCALTFHDEWIOYDYLQSLG 300
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
QY 361 HDNMTISIFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
DB 361 HDNMTISIFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
QY 421 LRVLVNDRVPLHGCANDKLGRCRDRDVEGLSFARSGGNWAECEFA 467
DB 421 LRVLVNDRVPLHGCANDKLGRCRDRDVEGLSFARSGGNWAECEFA 467

RESULT 13
US-09-155-855-1
; Sequence 1, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-1
Query Match 73.8%; Score 1818; DB 4; Length 443;
Best Local Similarity 75.7%; Pred. No. 4.8e-182;
Matches 334; Conservative 46; Mismatches 61; Indels 0; Gaps 0;
QY 27 NNSHCDTVDGGYOCPEISHLWGTSPYFSLADESAISDPDPCDRTVQVLSRHGARY 86
DB 3 NQSTCDTVDGGYOCFSETSHLWGTAPFESLANKSAISDPDPCDRTVQVLSRHGARY 62
QY 87 PTSSASKAYSALIEAIQKNATFAKGYAFLKTYNLTGADDTLPPFENOMVNSGKIFRYR 146
DB 63 PTSSASKAYSALIEAIQKNATFAKGYAFLKTYNLTGADDTLPPFENOMVNSGKIFRYR 122
QY 147 YKALARKIVPFTIRASGSDRVIASAEKFTGQSAKLADPGSQPHQASPVINIIPEGSGY 206
DB 123 YESLTRNIVPFTIRASGSDRVIASAEKFTGQSAKLADPGSQPHQASPVINIIPEGSGY 182
QY 207 NNTLDGCTCTAFEDSELGDDVDFANFTALFAPAIRARLEADLPVTLTDEVDVYLMDCPF 266
DB 183 NNTLDGCTCTAFEDSELGDDVDFANFTALFAPAIRARLEADLPVTLTDEVDVYLMDCPF 242
QY 267 DTVARTSDATELSPCALTFHDEWIOYDYLQSLGKYYGAGNPLGPAQGVGFANELIAR 326
DB 243 DTISTSTDTKLSFPFCDLFTHEEWINDYLSLNGKYYGAGNPLGPAQGVGFANELIAR 302
QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHNTMISIFALGLYNGTKPLSTTSV 386
DB 303 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHNTMISIFALGLYNGTKPLSTTSV 362
QY 387 ESHEETDGYASWTVPFAARAYVEMMQCAKEPVLVLRVNDVRVPLHGCANDKLGRCR 446
DB 363 ENITQTDGESSANTVFPASRMVEMMQCSEQLPLVLRVNDVRVPLHGCANDKLGRCR 422
QY 447 DDFVEGLSFARSGGNWAECEFA 467
DB 423 DDFVEGLSFARSGGNWAECEFA 443

RESULT 14
US-09-155-855-2
; Sequence 2, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 2
; LENGTH: 443
; TYPE: PROT
; ORGANISM: Aspergillus niger
; FEATURE:
; OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown
US-09-155-855-2

```

```

Query Match          73.6%; Score 1813; DB 4; Length 443;
Best Local Similarity 76.0%; Pred. No. 1.6e-181;
Matches 333; Conservative 46; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 30 SCDTVGGYQCPEISHLWGTYSFYPLADESAISPDVDDCRVTFVQVLSRHGARYPTS 89
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 6 TCDTVGGYQCPEISHLWGTYSFYPLADESAISPDVDDCRVTFVQVLSRHGARYPTS 65
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 90 SASKAYSALIEAIOKNATAFKGYAFLKTYNTLGGADDLTPFGENQMVNSGKIFRYRYKA 149
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 66 SKGKYSALIEAIOKNATAFKGYAFLKTYNTLGGADDLTPFGENQMVNSGKIFRYRYES 125
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 150 LARKIVPFRASGSRVITASAEKFTIEGQSAKLADPGSQPHQASPVINVIPEGSGYNT 209
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 126 LTRNIVPFRASGSRVITASAEKFTIEGQSAKLADPGSQPHQASPVINVIPEGSGYNT 185
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 210 LDHGTCTAFEDSELGDDVDFANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCPFDTV 269
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 186 LDPGCTVFEDSELADDIENATFATFVSIRORLENDLSGVSLLTDETVYVLMDCSFDTI 245
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 270 ARTSDATELSPCALFTHDEWIQDYQLSGKYYGYGAGNPLGPAQGVGFANELIARLTH 329
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 246 STSTVDTKLSPCCDLFTHEEWINYDQLSKNYYGHGAGNPLGPTQGVGYANELIARLTH 305
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 330 SPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMISIFFALGLYNGTKPLSTSVESI 389
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 306 SPVHDTSTNHTLDSNPATFPLNATLYADFSDHNTMISIFFALGLYNGTKPLSTSVESI 365
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 390 EETDGYASWTVPFAARYVEMMQCOAEKPELVRVLVNDVRVPLHGCADVCKLGRCKRDF 449
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 366 TQTDGFSSANTVPFAARYVEMMQCOAEKPELVRVLVNDVRVPLHGCADVCKLGRCKRDF 449
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 450 VEGLSFARSGGNWAECPA 467
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 426 VKGLSFARSGGDMGECFA 443
   :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```

RESULT 15

```

US-08-819-825-2
; Sequence 2, Application US/08819825
; Patent No. 5866118
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-819-825-2

```

```

Query Match          54.2%; Score 1334; DB 2; Length 475;
Best Local Similarity 56.1%; Pred. No. 3e-131;
Matches 273; Conservative 62; Mismatches 114; Indels 38; Gaps 11;
QY 1 MGVE-VVLLSTATLFGSTSGTALGP--RGNHSCDVTVDGGYQCFPEISHLWGTYSFYFSL 57
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 6 LGSELVLLQLQFSALL--TASPAIPPFWRKKHPNV-----IARHWGQYSPFSL 52
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 58 ADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFLK 117
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 53 AEVSEISPAVPKCRVEFFVQVLSRHGARYPTAKSEVYAEELLQRIQTATFEFGDAFLR 112
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 118 TYNTLGGADDLTPFGENQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKIEGF 177
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 113 DYAYHLGADNLTRFGEQMMESGRQFVHYREQAREIVPFVRAAGSARVIAAEFFNRGF 172
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 178 QSAKLADPGSQPHQASPVINVIPEGSGYNTLHDGTCCTAFEDSELGDDVE-ANFTALFA 236
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 173 QDAKDRDPRSNKQDAEPVINVIISETGNSNTLDGLTCPAAE--EAPDPTQPAEFLQVFG 230
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 237 PAIRARLEADLPVTLTDEDVYVLMDCPFDTVARTSDAT---ELSPFCALFTHDEWIQ 292
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 231 PRVLKTKTHMPGVNLTLEDVPLFMDLCPFDTVG--SDVLPFRQLSPFCHLTADDWMA 288
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 293 YDYLQSLGKYYGYGAGNPLGPAQGVGFANELIARLT-HSPVQDHTSTNHTLDSNPATFPL 351
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 289 YDYTYTLDKYYSHGGSGAFGSPRGVGFVNELIARMTGNLPVKDHTTVNHTLDDNPETPL 348
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 352 NATLYADFSDHNTMISIFFALGLYNGTKPLSTSV--SIEETDGYASASWTVPFAARYV 409
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 349 DAVLYADFSDHNTMTGIFSAMGLYNGTKPLSTSKIQPTGAAADGYAASWTVPFAARYV 408
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 410 EMMQC-----OAEKEPLVRVLVNDVRVPLHGCADVCKLGRCKRDFVEGLSFARSG 459
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 409 ELLRCETETSSSEEEGEDEPEFVRVLVNDVRVPLHGCADVCKLGRCKRDFVEGLSFARSG 468
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 460 GNWAECPA 466
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 469 GHWDRCF 475
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Search completed: October 26, 2001, 16:40:47
Job time: 4958 sec

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	1317.6	93.8	1426	9	AX021809	AX021809 Sequence	
2	783.6	55.8	1350	56	AF295325	AF295325 Synthetic	
3	614.2	43.7	1404	9	AI9452	AI9452 phytase cDN	
4	614.2	43.7	1404	10	I13430	I13430 Sequence 33	
5	614.2	43.7	1404	10	I33881	I33881 Sequence 19	
6	610.2	43.5	2000	13	ANPHYAG		
7	610.2	43.5	2363	9	AR018076	Z16144 A.niger phy	
8	610.2	43.5	2363	9	AR051916	AR018076 Sequence AR051916 Sequence	


```

/prim_transcript 166..1871
/partial
/feature="phyA"
/feature="phyA"
/function="primary transcript of phyA gene"
join(210..253,356..365)
/feature="phyA"
join(210..253,356..1715)
/feature="phyA"
/codon_start=1
/protein_id="CAA78904.1"
/db_xref="GI:2393"
/db_xref="SWISS-PROT:P34752"
translation="MGVSAYLLPLYLLSGVTSGLAVPASRNQSCDVTDOGYCFSET
SHLWQVAPFSLANESVISPVPAGCRVTEAOVLSHGARYPTDSKGYKYSALIEI
QONATFDGKAYFLKTYNSIGADDLTPFGEQELVNSGIFQRYBSLTNRNIPFIRS
SGSRVTSASKKFFIEGQSTKLKDPRAQPGSSPKIDVISEASSNNLDPDCTVFP
EDSLADTVEANFTATVPSPRIORLENDLSVLTIDTFTYLMDMG-SFDTISTVDIT
KLSPFDLFTHEWINDYDYLQSLKYYKHGAGNPLGPTGVGYANELIARLTHSPVHD
DTSSNHTLDSPPATFPLNSTLYADFSDNGIIISILFALGLYNGTKPLSTTVENITQT
DGFSSATVPFASRLYVEMMQCAEQEPLRVLVNDRVPLHSCPVDALGRCTRDSFV
RGLSFARSGDWAECFA"
254..355
/feature="phyA"
/number=1
356..1878
/number=2
369..1712
/feature="phyA"
1732..1760
/feature="phyA"

```

BASE COUNT 463 a 603 c 456 g 478 t

Query Match 43.5%; Score 610.2; DB 13; Length 2000;
 Best Local Similarity 64.9%; Pred. No. 4e-152;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

```

Qy 13 gtcgtgactgtccattgccacctgttctgggttccacatcccggtaccgcttgggtcct 72
Db 324 GTGGACTACTGATCGGTGACTATCTGTGCAGAGATCACTCCCGACTGGCGAGTCCCGCC 383
Qy 73 cgtgtaactctactctgtgacactgttgacggtgtgttaccagtgttccacagaatt 132
Db 384 TCGAGAAATCAATCCACTTCGATACGTCGATCAGGGGTATCAATGCTTCTCCGAGACT 443
Qy 133 tctcaactgtgggtacatactctccatcttcttcttgggtgacgaattctgtattct 192
Db 444 TCGCATCTTTGGGGTCAATACGACCCGTTCTCTCTGCGCAACGAATCGGTCTATCTCC 503
Qy 193 ccagacgttccaaagggtgttagagttacttctgttccagtttctgtagaacggtgt 252
Db 504 CCTGAGTGGCCGCGGATGCAGAGTCACTTCTGCTCAGGTCTCTCCCGTCTATGGAGCG 563
Qy 253 agatacccaacttctctgctgcttaagggtactctgcttctgtagaacgttcttcaaa 312
Db 564 CGGTATCCGACCGACTCCAGGCGCAAGAAATATCTCGCTCAGTCTCTCCCGTCTATGGAG 623
Qy 313 aacgtactgcttcaagggttagtagctgttcttctgtagaacgttcttcaaacgttgggt 372
Db 624 AACCGGACCACTTTGACGGGAAATATGCTCTCTGAGAGATACAACTACAGCTTGGGT 683
Qy 373 gctgacgactgactcattcgtgtagaacgaataagggttaactctgttattagttctac 432
Db 684 GCAGATGACTGACTCCCTTCGGAGAACAGAGAGCTAGTCAACTCCGGCATCAAGTCTAC 743
Qy 433 agaagatacaaggcttgggttagaagattgttccattcattagaccttctggttctgac 492
Db 744 CACCGGTACGAATCCGTCACAGGACATCTGTTCCATTCCGATCCCTCTGGCTCCACG 803
Qy 493 agagtattgcttctgtagaagttcattgaaggtttccaatctgtaagttggtgac 552

```

```

Db 804 CGCGTGTATCGCTCCGCCAAGAAATTCATCAGAGGGCTTCACAGAGCAACCAAGCTGAAGGAT 863
Qy 553 ccaggtgcttaaccacacacacagcttctccagttatttaagcttatttccagaagtgct 612
Db 864 CTTGCTGCCAGCCCGGCCAATCGTCCGCCAAGATCGACGTGTCTATTCCGAGGCCAGC 923
Qy 613 ggttacaacaacacttggaccacggttgggtgactgcttccgaaagaaactctggaatgggt 672
Db 924 TCATCCAAACACACTCTCCGACCAGCACCTGCACGTCTCTTCGAAGACAGCGAATTGGCC 983
Qy 673 gacgagtgaagtaacttcaactgctgttttctgctccacaaattagactagattggaa 732
Db 984 GATACCGTCGACGCCAATTCACCGCCACGTCGTGCTCCCTCCATTCGTCAACGCTGAG 1043
Qy 733 gctcacttgcaggtgttaacttgaactgacgacgacggttgaacttgaactgacacgtgt 792
Db 1044 AACGACGTGTCGGGTGTGACTCTCAGACACAGAAAGTACCTACCTCATGACATGTGC 1103
Qy 793 caattgacactgttctgtagaaacttctgacgactaacttctcctcattctgtgactgt 852
Db 1104 TCCTTCGACACCACTCTCCACGACCGTCGACACCAAGCTGTCCCTCTCTGTGACCTG 1163
Qy 853 ttcaactcagcagaatggattcaatacgaactacttgaacttcttgggttaagtactacgt 912
Db 1164 TTCACCCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTGAAAAGTATTACGGC 1223
Qy 913 tacggtgctgttaaccatttgggtccagctcaaggtgtgtgttctggttaacgaattgatt 972
Db 1224 CATGGTGCAGGTAACCGCGTCGCGCCGACCCAGGGGTGCGGTACGCTAACGAGCTCATC 1283
Qy 973 gctagattgactcaactctcagttcaagaccacacttcttactaaacacacttggactct 1032
Db 1284 GCCCGTCTGACCCACTCGCTGTCCACGATGACACCAAGTCCCAACACACACTTTGGACTCG 1343
Qy 1033 aacccagctacttctccattgaacgtacttctgtagcgtgacttctctcaacgaacact 1092
Db 1344 AGCCCGGTCTACCTTCCCGTCAACTCTACTCTACGCGGACTTTTCGCTATGACACGCGC 1403
Qy 1093 atggttctcttcttctggttgggttgggttgggttgggttgggttgggttgggttgggt 1152
Db 1404 ATCATCTCCATCTCTTGTCTTTAGGTCTGTGTACACAGGCACTTAAGCCGCTATCTACCAG 1463
Qy 1153 tctgttgaactctattgaagaaactgacggttactctgcttcttggactgttccattcgt 1212
Db 1464 ACCGTGGAGAAATATCACCCAGACAGATGATCTCTGCTGTGACGCTTCCGCTTCTCT 1523
Qy 1213 gctagagcttactggttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1272
Db 1524 TCGCGTTTGTACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
Qy 1273 ttggttaacgacagaggttcttccattgcaacggttgggttgggttgggttgggttgggt 1332
Db 1584 TTGGTTAATGATCGGTGTGTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1643
Qy 1333 aagagagacgacttctgtagaaggttcttctgtagatctggttgggttaacttgggaagaa 1392
Db 1644 ACCCGGGATAGCTTGTGAGGGGGTTGAGCTTGTGATAGTCTGCGGGTGTGCGGGAG 1703
Qy 1393 ttttgcctta 1403
Db 1704 TGTTTTGTCTTA 1714

```

RESULT 7

AR018076
 LOCUS AR018076 2363 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 7 from patent US 5780292.
 ACCESSION AR018076
 VERSION AR018076.1 GI:3973679
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

[illegible]

Db 578 TCGAGAAATCAATCCACTTGCATACGGTCGATCAAGGTATCAATGCTTCCGAGACT 637
QY 133 tctcactgtggtgacataacttccatttctctcttctggtgacgaatctgctattct 192
Db 638 TCGCATCTTTGGGGTCAATACGCGCGTCTCTCTCTGCGAAAGAAATCGGCATCTCC 697
QY 193 ccagacgttccaaaggtgttagagttacttctcgttcaagtttctctagacacggtgct 252
Db 698 CCTGATGTGCCCGCGGTTCAGAGTCACTTTCGCTCAGGTCTCTCTCCGTCATGGAGCG 757
QY 253 agatacccaacttctctgcttctaaagcgctactctctgcttctgattgaagctattcaaaag 312
Db 758 CGGTATCCGACCGAGTCCAAAGGGAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG 817
QY 313 aacgctactgctttcaagggtaagtaagcttcttgaagacttaacactcaactttgggt 372
Db 818 AACGTGACCACTTTGATGGAATAATATGCCCTCTCTGAAGACATACAACCTACAGCTTGGT 877
QY 373 gctgacgactgacttccattcgttgaacacaaatggttaactctggtatttaagttctac 432
Db 878 GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937
QY 433 agaagatacaaggttctgcttagaagattgttccattcatttagagcttctggttctgac 492
Db 938 CAGCGATACGAATCGCTCACAGGAACATCATTCCTTCATCCGATCCTCTGGCTCCAGC 997
QY 493 agagtattgcttctgctgaaggttcaattgaagtttccaaatctgtaagttggtgac 552
Db 998 CGGTGATGCGCTCGGGGAGAAATTCATTGAGGGCTTCCAGAGCACCAAGCTGAAGAT 1057
QY 553 ccaggtgtaaccacacacacagcttctccagttatttaacgttatttccagaaggtgct 612
Db 1058 CCTGCTGCCAGCGCGGCAATCTGTCGCCCAAGATCAGCTGTCATTTCCGAGGCCAGC 1117
QY 613 gttacacacacacttgggacacaggttctgctactgcttccagaatctgaattgggt 672
Db 1118 TCATCCAAACACTCTGACCCAGGACCTGTCACCTCTTTTGAAGACAGCAATTTGGCC 1177
QY 673 gacgactgtgaagtaacttcaactgctgttctgctccacacaaatagagctagattgaa 732
Db 1178 GATACCGTCGAAGCAATTTTACCGCCACGCTTCCGCCCTTCCATTCGTAACGCTGGAG 1237
QY 733 gctcaattgcaaggttgaacttgaactgacgaagctgttgaacttgaactggtgct 792
Db 1238 AACGACCTGTCTGGCGTGACTCTCACAGACAGAGTGAACCTACCTCATGGACATGTGC 1297
QY 793 ccattcagacactgttctagaacttctgaagctactcaattgtctccattctgtgacttg 852
Db 1298 TCCTTCAGACCACTCTCCACGACGCGTCGACACCAAGCTGTCCGCCCTTCTGTGACCTG 1357
QY 853 ttcactcagcagcaatggttcaatacagcacttctgcaatcttcttgggttaagtactacgt 912
Db 1358 TTCACCCATGACGAATGGATCCACTAGGACTACTCCAGTCCCTGAAAATACTACCGC 1417
QY 913 tacggtgctggaacccattggtccagctcaaggtgttgggttctgaacgaattgatt 972
Db 1418 CATGGCGAGGTAAACCGCTGCGCGGACCCAGGGGTCGGGTACGCTAACGAGCTCATC 1477
QY 973 gctagattgactcactctccagttcaagaccacacttctactaaccacacacttggactct 1032
Db 1478 GCCCGTCAACCACTTCGCTGTCCACGATGACACCGACTCCAAACACACACCTTGGACTCG 1537
QY 1033 aacccagctacttccattgaacgctacttctgacgctgacttctctcagcacaact 1092
Db 1538 AACCCAGCTACCTTCCCGCTCAACTCTACTCTACGCGGACTTTTCCACGATACCGC 1597
QY 1093 atggtttctatttctgcttgggttgaacagcgtaactaagccattgtctactact 1152
Db 1598 ATCATCTCATCTCTTCTTGGTCTGTATACAGGCACTTAAGCGGCTGTCTACCGC 1657
QY 1153 tctgttgaactctattgaagaaactgacggttactctgcttcttggactgttccattcgt 1212

Db 1658 ACCGTGAGAAATCAACCCAGACAGATGGCTTCTGCTTGGACGGTTCGGTTGCT 1717
QY 1213 gctagagcttaccttgaatgagtgcaatgtgaagctgaaagaaacacattggttagatt 1272
Db 1718 TCGCGTCTGTACGTCGAGATGATGCACTGCCAGCCGAGCAGGCGCTGGTGGCGTGC 1777
QY 1273 ttggttaacacagacagagttgttccattgacggttgggtgttgacaaagttgggttagatgt 1332
Db 1778 TTGTTTAATGATCGCTTGTCCCGCTGCATGGGTGTCCAATTGATGCTTTGGGAGATCT 1837
QY 1333 aagagagacgacttctgaaggttcttcttctgctagatctggttgtaactggaagaa 1392
Db 1838 ACCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTTGGGGGTGATTGGCGGAG 1897
QY 1393 tgtttcgctta 1403
Db 1898 TGTCTCTTGA 1908
RESULT 9
AR053934
LOCUS AR053934 2379 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5834286.
ACCESSION AR053934
VERSION AR053934.1 GI:5978796
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2379)
AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Fagerstrom,R.B.,
Miettinen-Oinonen,A.S.K., Turunen,M.K., Rambosk,J.A.,
Piddington,C.S., Houston,C.S. and Cantrell,M.A.
TITLE Recombinant cells that express phytate degrading enzymes in desired
ratios
JOURNAL Patent: US 5834286-A 1 10-NOV-1998;
FEATURES Location/Qualifiers
source 561 a 735 c 518 g 565 t
BASE COUNT 561 a 735 c 518 g 565 t
ORIGIN
Query Match 43.5%; Score 610.2; DB 9; Length 2379;
Best Local Similarity 64.9%; Pred. No. 4e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;
QY 13 gtcgtgctactgtccatgtccacacttgttcggttccacatccggtacogccttgggtct 72
Db 534 GNGGGACTACTGATCGCTGACAACTGTGAGAGTCACTCCGGACTGGCAGTCCCCGCC 593
QY 73 cgtggttaactctactctgtgacactgttgacggtgtgtaccatgtttcccaagaatt 132
Db 594 TCGAGAAATCAATCCACTTGCATACGGTCGATCAAGGGTATCAATGCTTCCGAGACT 653
QY 133 tctcactgtggtgataacttctccattctcttcttcttgggtgacgaatctgattct 192
Db 654 TCGCATCTTTGGGTCAATACGCGCGTCTCTCTCTGCGAAACGAATCGGCATCTCC 713
QY 193 ccagacgttccaaaggtgttagagttacttctcgaatttctgcttagacacggtgct 252
Db 714 CCTGATGTGCCCGCGGTTCGAGAGTCACTTTCGCTCAGGTCTCTCCGTCATGGAGCG 773
QY 253 agatacccaacttctctgcttctaaagcgctactctgcttcttgaagctattcaaaag 312
Db 774 CGGTATCCGACCGAGTCCAAAGGGAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG 833
QY 313 aacgctactgtttcaaggttaagtaagcttcttcttgaagacttaacactcaactttgggt 372
Db 834 AACGTGACCACTTTGATGGAATAATATGCCCTTCTCTGAAGACATACAACCTACAGCTTGGT 893
QY 373 gctgacgacttgaactcctcattcgttgaacacaaatggttgaactcgttattgaagttctac 432

Db 894 GCAGATGACCTGACTCCCTTCGGAGAGGAGGAGCTAGTCAACTCCGGCATCAAGTTCTTAC 953
 Qy 433 agaagatacaaggcttggctagaaaaattgttccattcattagagcttctgtgttgac 492
 Db 954 CAGCGATACGAATCGCTCACAAAGAACATCATTCGTTTCACTCCGATCCTCTGGCTCCAGC 1013
 Qy 493 agagtattgtctgtgtaaaagtcaattgaagtttccaattgtaagtgtgctgac 552
 Db 1014 CCGGTGATCGCTCCGGCGAGAAATTCATTGAGGGCTTCAGAGACCAACAGCTGAAGGAT 1073
 Qy 553 ccaggtgtaaccacacacacagcttctccagttattaaacgttattattccagaaagtgct 612
 Db 1074 CTTCTGTCACCGGGGCAATCGTCCGCCAAGATCGACGTGCTCAATTCGGAGGCCAGC 1133
 Qy 613 ggttaacaacacactttggaccaggttggctgtactgttctcgaagaatctgtaattgggt 672
 Db 1134 TCATCAACAACACTCTCGACCCAGGCACTGCACCTGTCTTTGAAGACAGCGAATTGGCC 1193
 Qy 673 gacgacgtgaagctaaactcactgctgttttcgctccaccaattagactagatgaa 732
 Db 1194 GATACCGTCGAAGCAATTCACCGCCACAGTTCGCCCTCCATTCGTCACAGCTCGAG 1253
 Qy 733 gctcacttgccaggtgttaacttgactgacgaagacgttggtaacttgatggacatggt 792
 Db 1254 AACGACCTGTCTGGCGTGACTCTCACAGACACAGAAAGTACCTCATGGACATGTC 1313
 Qy 793 ccattgacactgttgcctagaaactctgacgtactcaatgttccattctgtgacttg 852
 Db 1314 TCCTTCGACACCATCTCCACAGCACCGCTGCACACACCAAGCTGTCCCTCTGTGACCTG 1373
 Qy 853 ttacactcagcgaatgattcaatacactactgttgaacttcttgggttaagtactacgt 912
 Db 1374 TTCACCATGACGATGATGATCACTACGACTACCTCCAGTCCCTGAAAAAATACTACGGC 1433
 Qy 913 tacggtgtgttaaccacttgggtccagctcaggtgtgtgttctgttaacgaattgatt 972
 Db 1434 CATGGCGCAGGTAAACCGCTCGCGCCGACCCAGGCGCTGGCTAGCTAACGAGCTCATC 1493
 Qy 973 gctagattgaactcactctcagttcaagacacacttctactaacacacttggactct 1032
 Db 1494 GCCGCTCACCACACTCGCGCTGCTCCACGATGACACAGCTCCACACACACTTGGACTCG 1553
 Qy 1033 aaccagctacttccactgaacgtacttctgtacgtgacttctctcaagacacact 1092
 Db 1554 AACCAGCTACCTTCGCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAAACGGC 1613
 Qy 1093 atggtttctatttctcgttgggtttgttaacaaggtactaaagccattgtctactact 1152
 Db 1614 ATCATCTCTATCCTCTTCTTGGGTCTGTACAACGGCAGCTAAGCCGCTGTCTACACAG 1673
 Qy 1153 tctgtgaactctattgaagaaactgacgttactctgttcttggactgttccattcgt 1212
 Db 1674 ACCGTGGAGAAATATCACCAGACAGATGGGTCTCGTCTGCTTGGACGGTTCGGTTGCT 1733
 Qy 1213 gctagagcttaccttgaaatgatgcaatgtgaagctgaaagaaacccattgggttagatt 1272
 Db 1734 TCGGCTGTGACGTGAGATGATGACGAGTGCAGGCGGAGCAGGAGCGGCTGGCTGCTGTC 1793
 Qy 1273 ttggttaacgacagagatttccattgacaggttgggtgttgacaaagtgggttagatgt 1332
 Db 1794 TTGGTTAATGATCGCTTTCGCGCTGCAATGAGTGTGCTTGGGGGAGATGT 1853
 Qy 1333 aagagagacgacttctgaaagtttcttcttcgctagatctggttgtaactgggaagaa 1392
 Db 1854 ACCGGGATAGCTTTGTGAGGGGGTTGACGTTTGTCTAGATCTGGGGGTGATTGGCGGAG 1913
 Qy 1393 tgtttgcctta 1403
 Db 1914 TGTCTGCTTA 1924

RESULT 10

ASNPHTAS

LOCUS ASNPHTAS 2379 bp DNA PLN 23-MAR-1994
 DEFINITION Aspergillus niger var awamori phytase gene, complete cds.
 ACCESSION L02421
 VERSION L02421.1 GI:166518
 KEYWORDS phytase.
 SOURCE Aspergillus niger (strain ALK0243, sub-species awamori) DNA.
 ORGANISM Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE 1 (sites)
 AUTHORS Piddington,C.S., Houston,C.S., Paloheimo,M., Cantrell,M.,
 Miettinen-Oinonen,A., Nevalainen,H. and Rambosek,J.
 TITLE The cloning and sequencing of the genes encoding phytase (phy) and
 pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
 awamori
 JOURNAL Gene 133 (1), 55-62 (1993)
 MEDLINE 94040796
 REFERENCE 2 (bases 1 to 2379)
 AUTHORS Carter,J.R., Franden,M.A., Aebersold,R.H. and McHenry,C.S.
 TITLE Molecular cloning, sequencing and overexpression of the gene
 encoding the psi subunit of E. coli DNA polymerase III holoenzyme
 Unpublished (1992)
 JOURNAL Location/Qualifiers
 FEATURES
 source
 I. .2379
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 420..463
 /citation={2}
 /evidence=experimental
 join(420..463,566..1925)
 /EC_number="3.1.3.8"
 /citation={2}
 /codon_start=1
 /function="conversion of phytate to inositol and inorganic
 phosphorus"
 /evidence=experimental
 /product="phytase"
 /protein_id="AA16898.1"
 /db_xref="GI:166519"
 /translation="MGVSALLPLLYLAGVTSGLAVPASRNQSTCDTVDQVQCFSET
 SHLQGYAPFFSIANESALSDPVPACRVTFQVLSRHGARYPTESKGGKYSALLEEI
 QNNTTFDKGYAFLKTYNSLGDLLTPGEOLVNSGKIFQRYESLIRNIPFIRS
 SSSRVIASGEKFIQEGFQSTKLKDPRAOPGSSPKLDVVISEASSNNTLDPTCTVF
 ESELADTVEANFTATFAPSIQRLENLDSGLVTLDTETVYLMDCMSPDTISTSPVDT
 KLSPFCDLFDHEWIHYDLXLSKTYHGAGNPLGPTGQGVYANELARLHSPVHD
 DTFSSNTLDSNFPATPELTLADFSDHNGIISILFALGLYNGTLPSTTTVENTITQT
 DGFSSATVTFASRLYVENMQCAEQEPLRVLVNDRVPLVPLHGCPIDALGRCRTDSFV
 RGLSFARSGGDWAECSA"
 566..1925
 /note="putative"
 /citation={2}

BASE COUNT 561 a 735 c 518 g 565 t
 ORIGIN

Query Match 43.5%; Score 610.2; DB 13; Length 2379;
 Best Local Similarity 64.9%; Pred. No. 4e-152;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgctactgtccattgccacactgttgcggttccacatcccgtagccgcttgggtcct 72
 Db 534 GTGGGACTACTGATCGCTGACAAATCTGTGCAGAGTCACTCCGGACTGGCAGTCCCGGCC 593
 Qy 73 cgtggttaactctactcttctgacactgttgaogtgggttaccgaatttccccagaatt 132
 Db 594 TCGAGAAATCAATCACTTGGATACGTCGATCAAGGGTATCAATGCTTCTCGAGACT 653
 Qy 133 tctcaactgtgggttacatactctccattcttctctgtgactgaactgtctattct 192
 Db 654 TCGCATCTTTGGGGTCAATACGCCGCGTCTCTCTCTCGAAACGAATCGGCATCTCC 713
 Qy 193 ccagacgttccaaaggggtgttagagttacttctcgaagtctcgaatttcttagacacggtgct 252

QY 313 aacgctactgtttcaagggttaagtaagcgttttcttgaagacttaacactacacttttgggt 372
DB 624 AACGGACACCTTTGACGGAAATATGCTTCTCTGAAGACATACAACTACAGCTTGGT 683
QY 373 gctgacgactgactcattcgttggaacacaaatgttaactctgtatataagttctac 432
DB 684 GCAGATGACCTGACTTCCGAGAGACAGGAGCTAGTCAACTCCGCATCAAGTCTAC 743
QY 433 agaagatacaagggtttggttagaagaattgttccattcattagagcttctgttctgac 492
DB 744 CAGCGTACGAATCGTCACAAAGAAACATGTTCCATTCCGATCCTCTGGCTCCAGC 803
QY 493 agagtattgtctgtgtaaaagtcaatgaagtttccaaatctgctaagttgctgac 552
DB 804 CCGGTGATCGCTCCGCAAGAAATTCATCGAGGCTTCCAGAGCACCACCAAGCTGAAGAT 863
QY 553 ccaggtgtcaaccacacacagctctcagttattcaagttattattccagaaggtgct 612
DB 864 CTTGTCGCCAGCCCGGCAATGTCGCCCAAGATCGACGTGGTCTATTCGAGGCCAGC 923
QY 613 ggttaacacacactttggaccaggtttgtgtactgtcttcgaagaatctgaattgggt 672
DB 924 TCATCAACAAACACTCTCGACAGGACCTGCACTGTCTTCGAAGACAGCGAATGGCC 983
QY 673 gacgagttgaagctaaacttcaactgtgttttgcctccaaatagagctagattgaa 732
DB 984 GATACCGTCGAAGCAATTTACCGCCACGTTGCTGCCCTCCATTCGTCACGCTGGAG 1043
QY 733 gctcactgtcaggtgttaacttgactgacgaagcgtgtgttaacttgatgacatggt 792
DB 1044 AACGACCTGTCGGGTGACTCTCACACACACAGAGTGACCTACCTCATGACATGTGC 1103
QY 793 ccattcgacactgtgttagaactctgacgactcactcaatgtgtctcaattctgtgacttg 852
DB 1104 TCGCTTCGACACCACTCCACACGACCTCGACACCAACAGCTGTCCCTCTGTGAOCTG 1163
QY 853 ttcaactcagcagaatgaattcaatcacgactacttgcgaactttgttggaactacagt 912
DB 1164 TTCACCATGACGAATGATCACTACGACTACCTTCAGTCTCTTGAAAGATATACGGC 1223
QY 913 tacggtgtgttaacccattgggttcagcgtcaaggtgtgtgttggtttaacgaattgatt 972
DB 1224 CATGTGTCAGGTAAACCGCTCGCGCCGACCCAGGCGCTCGGCTACGCTAACGAGTCACT 1283
QY 973 gctagattgactcactctccattcaagaccacacttctactcaacacacactttgactct 1032
DB 1284 GCGGCTCAACCCACTCGGCTTCCACGATGACACCAAGTTCACCAACCACTTTGACTCG 1343
QY 1033 aacccagctacttccattgacgctactttgtacgctgacttctctcaagacaacact 1092
DB 1344 AGCCGGCTACCTTCCGCTCAACTCTACTCTCTACGGGAGCTTTTCGATGACAAACGGC 1403
QY 1093 atggtttctatttcttctgtgtgtttgttgaacaggtactaagccattgtctactact 1152
DB 1404 ATCATCTCAATCTCTTCTTCTAGTCTGTACAAACGCGACTAACGCGCTAFTACACAG 1463
QY 1153 tctgttgaactcttgaagaactgacgcttactctgcttcttggactgttccattcgt 1212
DB 1464 ACCGTGGAGAAATACCCACAGACAGATGATCTCTGCTGTGACGCTTCGCTTGGT 1523
QY 1213 gctagagcttactgtgaatgatgcaatgtgaagctgaaaggaaccattgtgttagatt 1272
DB 1524 TCGGTTTGTACGTCGAGATGATGATGTCAGTGTACGGGGAGCAGGAGCGCTGTGCTGTC 1583
QY 1273 ttggttaacagagaattgttccattgacggtgtgtgtgtgtgacaggttgggttagatt 1332
DB 1584 TTGGTTAATGATCGCGTGTGTCGCTGTGATGGGTGTCCGCTGTGATGCTTTGGGGAGATGT 1643
QY 1333 aagagagagacttctgtgaaggtttgttcttctgttagatctgtgtgttaacttgggaagaa 1392
DB 1644 ACCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTTAGATCTGGGGGTGATTTGGCGGAG 1703
QY 1393 tgtttcgctta 1403

DB 1704 TGTTTTGTCTTA 1714
||||| |||||
RESULT 14
AX000634
LOCUS AX000634 1571 bp DNA PAT 10-MAR-2000
DEFINITION Sequence 5 from Patent EP0897010.
ACCESSION AX000634
VERSION AX000634.1 GI:7241015
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1571)
AUTHORS Van, L.A. and Wyss, M.
TITLE Modified phytases
JOURNAL Patent: EP 0897010-A 5 17-FEB-1999;
HOFFMANN LA ROCHE (CH)
FEATURES
Location/Qualifiers
1..1571
/organism="unidentified"
/db_xref="taxon:32644"
Join(43..89,147..1497)
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB77103.1"
/db_xref="GI:7241016"
/translation="MVTLTFLSAAVLLSGBVSAAPSSAGSKSCDITVDLGYQCSPATS
HLWGYSPPFSLSELSVSKLPDKCRITLVLSRHGARIPTSSKKYKLVTAIQ
ANATDKGFAPLKYNTYLTGLADLTPEEQOLVNSGIFKYORKALARSVVPFIRAS
GSDRVIASGEKFIETFOQAKLADPGATNRAAPASVITPESETFNNTLDHGVCTKFEA
SOLGDEVAANTFALFAPDTRARAEKHLPGVTLTDEDVSLMDMCSFDTVAFTSDASOL
SPCOLPTHNEWKYVQLQSLKYYGAGNPLGPAOGIGFTNGLIARLTSPVQDHT
STNSLVSNPATFPLNATWYDFSHDNSMVSEFALGLYNGTEPLSRTSVPSAKELDG
YASWVVPGARAYFETWCKSEKEPLVRLINDRVPLHGCDDVDKLGRCRLNDFVKS
LSWASGGNMGWCF"
BASE COUNT 326 a 416 c 453 g 376 t
ORIGIN
Query Match 43.4%; Score 610; DB 9; Length 1571;
Best Local Similarity 66.7%; Pred. No. 4,4e-152;
Matches 887; Conservative 0; Mismatches 440; Indels 3; Gaps 1;
QY 75 tggtaactctcaactcttggacactgttgacggtgtgttaccactgtttccacagaaatttc 134
DB 171 TGTGGCTCCAAGTCTCGATACGGTAGACCTCGGTACCGTCTCCCTCGGACTTC 230
QY 135 tcaactgtgggtgtacatactctcattctcttcttggctgacgaatctgctattctctcc 194
DB 231 TCATCTATGGGGCAGTACTCGCATTTCTTCGCTCGAGGAGAGCTGTCCCTGTGCGAG 290
QY 195 agacgttccaaagggtttagagttactcttgcgttcaagttttgtctagacacggtgctag 254
DB 291 TAAGCTTCCCAAGGATTTCGGGATCACCCTGGTGTGATCGCGCATCGAGGGCG 350
QY 255 ataccacacttctctcgtcgtcgaagcgtactctgcttggattgaagctattccaaagaa 314
DB 351 GTACCAACAGCTCCAAGAGCAAAAAAGTATAAGAAAGCTTGTGACGGCATCCAGGCCAA 410
QY 315 cgtactgtttcaagggttaagtaagcgtttcttgaagacttaacactacactttgggtgc 374
DB 411 TGCACCCGACTTCAAGGGCAAGTTTGCCTTTTGAAGACGCTACAACTATACTCTGGGTGC 470
QY 375 tgacgactgactcattcgttggaacacaaatgtttaactctggtatttaagttctacag 434
DB 471 GGATGACCTCACTCCCTTGGGGAGCAGCAGCTGTGTAACCTCGGGCATCAAGTCTACCA 530
QY 435 aagatacaaggcttggcttagaagattgttccattcattagagcttctgttctgacag 494
DB 531 GAGGTACAAGGCTCTGGCGCGCAGTGTGGTGGCTTTATTTCGCGCCTCCTGAGGCTCGGACCG 590

QY	495	agttattgctctgctgaagaagttaattgaagggttttccaattctgctaagttgggtgaccc	554
Db	591	GGTATTGTCTCGGGAGAGAAGATTATCAGGGGTTCCAGCAGCGGAAGCTGGCTGATCC	650
QY	555	aggTgctaaccacacacaagattctccagttatttaaacgigtattattccagaagtgctb99	614
Db	651	TGCGCGGA --- CGAACGGCGCGCTCCGGCGATTAGTGTGATTATTCGGAGACCGGAGAC	707
QY	615	ttacaacaacactttggaccacggtttgtgactgcttccgaagaactgaattgggtga	674
Db	708	GTTCAACAATACGCTGGACACGGTGTGTGCACGNAAGTTTGAGCGGAGTCACTGGGAGA	767
QY	675	cgagtgtaagctaacttaacttaactgctggttttcgctccaccaattagactagattggaagc	734
Db	768	TGAGGTGTGGGGCAATTTCACTTCGCCTCTTTGCACCCGACATCCGAGCTCGCGCGGAA	827
QY	735	tcacttgccagggtttaacttgactgacgaagacgttgtaacttgatggacatgctcc	794
Db	828	GCATCTTCCTGGCGTGACGCTGACAGACGAGCGGTGTCAAGTCTAATGACACATGTGTTC	887
QY	795	attcgacactgttgtagaactctgacgttactcaattgctctccattctgtgactggtt	854
Db	888	GTTTGATACGGTAGCGCGCACCGGACGCAAGTCAGCTGTCCCGTCTGTCTCACTCTT	947
QY	855	cactcaagacgaatggattccaatacagactacttgcaactctttgggtaagtactacggtta	914
Db	948	CACTCAATGAGTGGAAAGAGTACAATCACTTCACTCTTGGCGAAGTACTACGGCTA	1007
QY	915	c9gtgctgtaaccattgggtccagctccaaggltggtttcgttaacgaattgattgc	974
Db	1008	CGSGCAGCAACCCCTCTGGACCGGCTACGGGATAGGGTTCAACCAACGAGTGATGTC	1067
QY	975	tagattgactactctccagttccaagacacacttctactaaccacactttggactctaa	1034
Db	1068	CGGGTTGACTCTGTCGCGAGTGCAGGACACACCAGCACTAACTCGACTCTAGTCTCCAA	1127
QY	1035	ccagctactttccattgaagcactttgtgaagctgactctctccacgacacactat	1094
Db	1128	CCCGGCCACTCTCCCGTTGAACGCTACCATGTAGCTGCGACTTTTTCACACGACACAGCAT	1187
QY	1095	ggtttctattttctcgctttgggtttgtacaacggtactaaagcattgtctactacttc	1154
Db	1188	GGTTTCCATCTCTTTGATTTGGCCCTGTACACGGCAGCTGAACCCCTGTCCCGGACCTC	1247
QY	1155	tgttgaatctattgaagaactgacggttactctgctcttggactgttccattcctgcgc	1214
Db	1248	GGTGAAAACGCCAAGAAATTTGGATGGGTATTCTGTCATCTCGGTGGTGGCTTTCCGGCGC	1307
QY	1215	tagagcttactgttaaatgatgaatgtgaagctgaaagaaaccattggttagatttt	1274
Db	1308	CGGACGCTACTCTCGAGACGATGCAATGCAAGTCCGAAAGAGCGCTCTTTTCGCGCTTT	1367
QY	1275	ggttaacgacagagtgtgtccattgcacggtgtg9gttgtagaagttgggtagatgtaa	1334
Db	1368	GATTAATGACCGGGTTGTGCCACTGCAATGCTGGGATGTGGACAAAGTGGGGCGATGCA	1427
QY	1335	gagagacactctgtgaaggtttgtcttcgctagatctggttggttaactgggaaagt	1394
Db	1428	GCTGAATGACTTTGTCAAGGGAATAGTGTGGCCAGATCTGGGGCACTCTGGGGAGGTG	1487
QY	1395	tttcgcttaa	1404
Db	1488	CTTAGTTGA	1497

RESULT **15**

SFU59804	.	DNA	PLN	17-JAN-1998
LOCUS	:	1612 bp		
DEFINITION		<i>Aspergillus fumigatus</i> pyrtase gene, complete cds.		
ACCESSION		U59804		
VERSION		U59804.1	GI:2108353	
KEYWORDS		. Aspergillus fumigatus.		
SOURCE				

ORGANISM	Aspergillus fumigatus					
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;					
AUTHORS	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.					
TITLE	Pasamontes L., Haiker M., Wyss M., Tessier M. and van Loon A.P.					
JOURNAL	Phyase cloning, purification, and characterization of a heat-stable					
MEDLINE	Appl. Environ. Microbiol. 63 (5), 1696-1700 (1997)					
REFERENCE	97288063					
AUTHORS	Pasamontes L.					
TITLE	Direct Submision					
JOURNAL	Submitted (02-JUN-1996) Luis Pasamontes, VFCB, F.Hoffmann-La Roche					
FEATURES	Ag., Basel 4070, Switzerland					
source	Location/Qualifiers					
	1..1812					
	/organism="Aspergillus fumigatus"					
	/db_xref="taxon:5085"					
TATA_signal	195..198					
TATA_signal	197..200					
sig_peptide	join(284)..330,388..418)					
mRNA	join(<284..330,388..>1738)					
CDS	join(284)..330,388..1738)					
	/EC_number="3.1.3.8"					
	/codon_start=1					
	/product="phytase"					
	/protein_id="AAB96872.1"					
	/db_xref="GI:2108354"					
	/translation="MVTITFLLSAAIYLLSGRVSAASPAGSKSCDVTDLGOCSPATS					
	HLWGQFPFSELDLSSVKKPKDRITLVQLSRHGARYPTSSKSKKYKLIVTAIO					
	ANATDFKFAPFKTYNTLGADLTLPGEOLNSRGIKFYORYKALARSVPVFIRAS					
	GSDRYVASEGKEFIQGQAOKADPGATNRAIPAISVIIPESETFNNTDHGVCTNFEA					
	SQLGDVEANTALPAPIRARAEXHLEFGVLITDEDVSLMDMSFDIVARTSDASQL					
	SFFCOLFHNEKKNTYLQSUGKIYGYGAGNPGLPAQGIETNELIAELRTSPVOODHT					
	STNSVLVSNPATFFPLNATMYVDIFSNDNSWVSIFFAVLNGYTEPLRSITSVPSAKELD					
	YSASWVPFGARAYPETMQCKSEKEPLVRALINDRVPLHGGDVDKLRCKKLINDEVKG					
	LSWAARGGNWGECFS"					
BASE COUNT	377 a	480 c	532 g	423 t		
ORIGIN						
Query Match	43.4%; Score 610; DB 15; Length 1812;					
Best Local Similarity	66.7%; Pred. No. 4.5e-152;					
Matches 887; Conservative	0; Mismatches 440; Indels 3; Gaps 1;					
QY	75	tggtaactcactcttgtagacactgttgacagdggtgtaccgaattttccccaagaatttc	134			
Db	412	TGCTGGCTCCAAGTCTCGCATACGGTAGCGGTACCACTCGGGTACCAGTCGCCCTGCACATT	471			
QY	135	tcacttgtgggtgcatactctccatctctctcttggtgcagcaatctgcgtattctccc	194			
Db	472	TCATCTATGGGGCCAGTACTCGCCATTCTTTTCGTCGAGGACGAGCTGTCGCTGTCCGAG	531			
QY	195	agcgctccaaagggttagagtaccttcttcgaagtgttgctagacacggtgtcgtag	254			
Ddb	532	TAAGCTTCCCAAAGGATTCGCCGATCACCTTGTCAGGTGCTATCGCGCATTGGAGCGCG	591			
QY	255	ataccacaactcttcggtctaaggcgtactctgtcttgattgaagctattcacaaga	314			
Ddb	592	GTACCCAACCACTCCAAAGACAAAAGATTAAGAAGCTGTGACGGCGATCCAGGCCAA	651			
QY	315	cgtactgtttcaaagggtgaagtagcgtctcttgaagacttacactacactttgggtgc	374			
Db	652	TGCCACCGACTCAAGGCCAAGTTTGCCTTTTGAAGACGTACAATACTCTGGGTGC	711			
QY	375	tgaqgacttgactcatcgttgtagaacacaaaatggttaactctggtattaatgtctacag	434			
Ddb	712	GGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGGTGAACTCGGGCGATCAAGTCTACCA	771			
QY	435	aagatacaagcgtttggcataaagattgttccattcattagactcttggttctgacag	494			
Db	772	GAGGTACAAGCTCTGGCCCCCAGTGTGGTGGCGGTTTTATTCCGCGCTCAGCTCGGACCG	831			

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*
48: gb_est48:*
49: gb_est49:*
50: gb_est50:*
51: gb_est51:*
52: gb_est52:*
53: gb_est53:*
54: gb_est54:*
55: gb_est55:*
56: gb_est56:*
57: gb_est57:*
58: gb_est58:*
59: gb_est59:*
60: gb_est60:*
61: gb_est61:*
62: gb_est62:*
63: gb_est63:*
64: gb_est64:*
65: gb_est65:*
66: gb_est66:*
67: gb_est67:*
68: gb_est68:*
69: gb_est69:*
70: gb_est70:*
71: gb_est71:*
72: gb_est72:*
73: gb_est73:*
74: gb_est74:*
75: gb_est75:*
76: gb_est76:*
77: gb_est77:*
78: gb_est78:*
79: gb_est79:*
80: gb_est80:*
81: gb_est81:*
82: gb_est82:*
83: gb_est83:*
84: gb_est84:*
85: gb_est85:*
86: gb_est86:*
87: gb_est87:*
88: gb_est88:*
89: gb_est89:*
90: gb_est90:*
91: gb_est91:*
92: gb_est92:*
93: gb_est93:*
94: gb_est94:*
95: gb_est95:*
96: gb_est96:*
97: gb_est97:*
98: gb_est98:*
99: gb_est99:*
100: gb_est100:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

Db 140 GATGTTTGTGAAAAATGACTTGGCAGGCGAGAACGAGGAGCTTGAGAACTCTGGT 81
Qy 1278 taacacagagttgttcacgtgacggttggtgttgacaaagttagatgaag 1337
Db 80 CAACGACAGGGTGACGCGCTGCAAGACTGCGATGCCGACGATGGGTCTGTCACGCT 21
Qy 1338 agacgacttcgtgaaggtt 1357
Db 20 GAGCAAGTCTGTTGAGAGCT 1

RESULT 2
LOCUS A0255459 837 bp DNA GSS 23-OCT-1998
DEFINITION mgxb0014M15r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0014M15r, DNA sequence.
ACCESSION A0255459
VERSION A0255459.1 GI:3779774
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea.
REFERENCE 1 (bases 1 to 837)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 60
High quality sequence stop: 277.
Location/Qualifiers
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0014M15r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 248 a 154 c 188 g 243 t 4 others
ORIGIN

Query Match 8.4%; Score 117.8; DB 226; Length 837;
Best Local Similarity 56.2%; Pred. No. 2.2e-23;
Matches 221; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 737 acttgcaggtgttaacttgactgacgaagacgtgttgtaactgagcatgtgccat 796
Db 140 ATTTTCAAGGTGCTGGGTATACCGGTCAAGACGTTGTAAATCTATATGGGACTGTGACCA 199
Qy 797 tcgacactgtgtgacgaactctgacgctactcaattgtctccattctgtgactgttca 856

```

```

Db 200 TGGAAACACAGCGCAACTTTGAAAAAATGGCCAACTTCACCGCTTTCACATCTGTTTA 259
Qy 857 ctacacgaatgattcaatcacgactacttgcaactctttgggtgaagtactacggttacg 916
Db 260 CGAAGCAGATTGGGTAAATATGTTACTTGTCCAGCGTGCAAAAATGGTACAGATACG 319
Qy 917 gtgtgtgtaaccattgggtccagctccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 976
Db 320 GAAACGGTAATCTTTGGGCCCAACTATGCGGTGGGTGAAACGAACCTATTATGCAAC 379
Qy 977 gattgactcaactccagttcaagacacacacttctactaacacacttggactctaac 1036
Db 380 GATTAAACCCGAAGCCAGTCCTCAAGATCAACACCATGTCCAAATAGCACACTTGACATGAACC 439
Qy 1037 cagctacttctccactgaacgctactttgtacgctgactctctctcagcacaacactatgg 1096
Db 440 CGGNAACCTTACTCTTACAAAGCAAAATGTGTCTGATTTATCCATACGGAGGATATTA 499
Qy 1097 ttctattttctctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1129
Db 500 TAGGGATTACGCGCGCTTTGGGCTATTTAACG 532

RESULT 3
LOCUS AJ274007 581 bp mRNA EST 29-DEC-1999
DEFINITION AJ274007 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae
cDNA clone Ma#1628, mRNA sequence.
ACCESSION AJ274007
VERSION AJ274007.1 GI:6433380
KEYWORDS EST.
SOURCE Metarhizium anisopliae.
ORGANISM Metarhizium anisopliae.
REFERENCE 1 (bases 1 to 581)
AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.
TITLE EST analysis of the insect pathogenic fungus Metarhizium anisopliae
JOURNAL Unpublished (1999)
COMMENT Contact: Screen SE
Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA.
Location/Qualifiers
/organism="Metarhizium anisopliae"
/strain="ARSEF 2575"
/db_xref="taxon:5530"
/clone="Ma#1628"
/note="Vector: Unizap; Metarhizium anisopliae was grown on
insect cuticle for 24 hours. A cDNA library was
constructed in the unidirectional Lambda vector, Unizap"
BASE COUNT 138 a 153 c 146 g 144 t
ORIGIN

Query Match 7.8%; Score 109; DB 104; Length 581;
Best Local Similarity 57.8%; Pred. No. 7.4e-21;
Matches 234; Conservative 0; Mismatches 165; Indels 6; Gaps 2;

Qy 142 tgggtatcatctctccattctcttctgtgctgacgaactgctgtcttctccagagctt 201
Db 163 TGGGCCAGTACTCGCGTACTCTC---AGCACACAGGAGTCTATTATAGCCTGACATC 219
Qy 202 ccaagaggtgttagagttacttctgttcgaagttttgttagacacggtgtgtagatccca 261
Db 220 CCGTCGGGATGCGAGTCCACCTTTGCTTCAATCTGTCTCGACATGGGTCCCGCAACCCG 279
Qy 262 acttcttctgctcaagcgtactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 321
Db 280 ACGGCGCGCAAGTCCAGGCCCTACAAGGACCTGTGTGAGCGGATCCAAAGAGAGCTCAAG 339

```

QY 322 gcttcaagggttaagctacgcttcttgaagacttacactacacttctggtgctgacgac 381
 Db 340 GACTATGGGAAGGCTTCAGATTTCTCAAGATTAACAAGTACTACTAGGTTCTCAGGAC 399
 QY 382 ttgactccattcgttgaaacacaaatggttaactctggttattaaagtctacagaagatac 441
 Db 400 CTCACCCCTTGGCGGAGGAGGAATGTCAGTCTGGAAGAAAGTCTTCAAGGCTAT 459
 QY 442 aagccttggctagaa---gattgtccattcattagagcttctggttctgacagagtt 498
 Db 460 CAGAAGTACGGAGGATTCAGCCATCCCTTTGTTGAGCCTCGGGCTCTGAAAGATC 519
 QY 499 attgcttctgctgaagaagttcattgaaggtttccaatctgctaa 543
 Db 520 GTCATGTCAGCGCAAGATTTTGTTCATGCTTCTACAAAGCCAG 564

RESULT 4
 AQ163004 753 bp DNA GSS 09-SEP-1998
 LOCUS mgxb0021D19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 DEFINITION clone mgxb0021D19r, DNA sequence.
 ACCESSION AQ163004
 VERSION AQ163004.1 GI:3559405
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 753)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 448.
 Location/Qualifiers
 1..753
 /organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0021D19r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."
 208 a 180 c 156 g 206 t 3 others

Query Match 6.0%; Score 83.8; DB 225; Length 753;
 Best Local Similarity 50.7%; Pred. No. 2e-13;
 Matches 227; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

QY 90 ttgtgacactgttgacggtggtttaccatggtttccagaaatttctcacttctggtggtac 149
 Db 189 TTTTGATCGCCAGCCGCGGTTCATTTAAAGCAGCAATCACTCATATATTGGGCCA 248
 QY 150 ataactcattctctctcttctggtgacgaatctgctatttctccagacgttccaa---- 205
 Db 249 GTATGCACCAATTTTCGTTGCACCGTCAGCACCATCGAATATGATTCCTCTGTACCTT 308
 QY 206 --agggttgttagacttacttctcgttcaagttttgtctagacaggtgctagataccaac 263
 Db 309 GCCAGGTTGTAGCATTTAGCTTTGCCAAGTCTTCAGAGCAGCGGGCTCGATACCCAAC 368
 QY 264 ttcttctgctctaaagcgctactctgtgttgattgaagctattcaaaagaacgctactgc 323
 Db 369 CGCTCAAAACCGCGGGAATTTCTGAAACACGATTTCACCGACTCCAAACCTCACGTACCG 428
 QY 324 ttcaagggttaagtacgcttcttgaagacttacactacactttgggtgctgacgactt 383
 Db 429 TAGTGGCGTATTGAATTAACATCAATTAATAAATACTCAATTCGAGTCGAGGAATN 488
 QY 384 gactccattcgttgaaacacaaatggttaactctggttattaaagtctacagaagatacaa 443
 Db 489 GAATGATTTGGCGCTCGGCAACAGAGAACTCGGGTATTTATTTTACCAAGATACCA 548
 QY 444 ggcttggctagaagaattgtccattcattgagcttagagcttctggttctgacagagttattgc 503
 Db 549 AAATCTCGCCAGAGGAACGAGCACTTATTCGTCTATGATGACAAANAACCGCTGTTGA 608
 QY 504 ttctgctgaaagtctcattgaaagtttc 531
 Db 609 TAGCGCAGAACTCTGGGCTAGGGGTTTC 636

RESULT 5
 AQ162040/c
 LOCUS mgxb0010F06r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 DEFINITION clone mgxb0010F06r, DNA sequence.
 ACCESSION AQ162040
 VERSION AQ162040.1 GI:3558441
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 699)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 285.
 Location/Qualifiers
 1..699
 /organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0010F06r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."
 208 a 180 c 156 g 206 t 3 others

Query Match 6.0%; Score 83.8; DB 225; Length 753;
 Best Local Similarity 50.7%; Pred. No. 2e-13;
 Matches 227; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

BASE COUNT
ORIGIN

175 a 191 c 148 g 185 t
Query Match 5.8%; Score 81.4; DB 225; Length 699;
Best Local Similarity 59.9%; Pred. NO. 9.7e-13;
Matches 136; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1178 acggttactctgttctgactgttccattcgctgctagagcttcaatgaaatgatgc 1237
DB 651 ACCGGTTCGTCAGTTCGGCGGTACCGTTCGCAGCGAATGTTGTGAAATAATGA 592
QY 1238 aatgtgaagctgaagaagcaaccattggttagagttttggttaacgacagagtggttccat 1297
DB 591 CTTCGCAGGCGAGACGAGGAGCTTGTGAGAACTCCTGGTCAACGACAGGGGTGACGCGC 532
QY 1298 tgcacggttggtgtgacaagttggttagatgaagagagacacttcgttgaggtt 1357
DB 531 TGCAGAACTGCGATGCGACAGTATGGGTGCTGCACGCTGACGAGTTCGTTGAGAGCT 472
QY 1358 tgccttcgttagatctgtggttaactgggaagaatgttctccttaa 1404
DB 471 TAAGCTTCGAGAGTGGAGGTCGCTGGATCAATGTTTGTCTGA 425

RESULT 6

AQ161556/c
LOCUS AQ161556 810 bp DNA GSS 09-SEP-1998
DEFINITION mgxb0008K02r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0008K02r, DNA sequence.
ACCESSION AQ161556
VERSION AQ161556.1 GI:3557957
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
1 (bases 1 to 810)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 321.
Location/Qualifiers
1. .810
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with

JOURNAL

COMMENT

Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 321.
Location/Qualifiers
1. .810
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with

FEATURES
source

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

BASE COUNT
ORIGIN

196 a 219 c 188 g 207 t
Query Match 5.4%; Score 75.6; DB 225; Length 810;
Best Local Similarity 58.4%; Pred. NO. 5.1e-11;
Matches 132; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1179 cggttactctgtcttctgactgttccattcgctgctagagcttcaatgaaatgatgca 1238
DB 650 CCGGTTCTCGTCAGCTFGGCGGTACCGTTCGCGAGCGAGATGTTTGTGAAAAATGAC 591
QY 1239 atgtgaagctgaagaagcaaccattggttagagttttggttaacgacagagttgttccatt 1298
DB 590 TTGCGCAGGGCAGAACGAGGAGCTTGTGAGAACTCCTGGTCAACGACAGGGGTGACGCGCT 531
QY 1299 gcaagctgtgtgtgtgacaaagttgggttagatgtaagagagagagacttcgttgaagttt 1358
DB 530 GCAGAACTGCGATGCCCGACAGTATGGGTGCTGCACGCTGACGAAGTTCGTTGAGAGCTT 471
QY 1359 gtcttcgtagatctgtgtgtaactgggaagaatgttctcgttaa 1404
DB 470 AAGCTCTCGAGGAGTGGAGTTCGCTGGATCAATGTTTGTCTGA 425

RESULT 7

AQ361474
LOCUS AQ361474 769 bp DNA GSS 03-FEB-1999
DEFINITION mgxb0004I12f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0004I12f, DNA sequence.
ACCESSION AQ361474
VERSION AQ361474.1 GI:4211313
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
1 (bases 1 to 769)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 51
High quality sequence stop: 416.
Location/Qualifiers
1. .769
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII; Rice blast is one of the most devastating fungal diseases

JOURNAL

COMMENT

Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 51
High quality sequence stop: 416.
Location/Qualifiers
1. .769
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII; Rice blast is one of the most devastating fungal diseases

FEATURES
source

[illegible]


```

RESULT 11
AQ020816
LOCUS
DEFINITION
  AQ020816 481 bp DNA GSS 20-APR-1999
  RPCI11-48J14.TK RPCI-11 Homo sapiens genomic clone RPCI-11-48J14,
  DNA sequence.
ACCESSION
  AQ020816
VERSION
  AQ020816.1 GI:3613607
KEYWORDS
  GSS.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 481)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  Use of human BAC End Sequences for Sequence-Ready Map Building
  Unpublished (1998)
  Other_GSSs: RPCI11-48J14.TJ
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdamas@tigr.org
  Clones are derived from the human BAC library RPCI-11. For BAC
  library availability, please contact Pieter de Jong
  (pieter@dejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
  Research Genetics (info@resgen.com). BAC end search page:
  http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
  Class: BAC ends.
  Location/Qualifiers
    1..481
    /organism="Homo sapiens"
    /db_xref="GDB:751827"
    /db_xref="taxon:9606"
    /clone="RPCI-11-48J14"
    /clone_lib="RPCI-11"
    /sex="Male"
    /cell_type="Lymphocytes"
    /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
    RPCI11 Human Male BAC Library"
  BASE COUNT 150 a 99 c 81 g 151 t
  ORIGIN

  Query Match 3.0%; Score 41.8; DB 225; Length 481;
  Best Local Similarity 45.5%; Pred. No. 0.35;
  Matches 148; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

  QY 966 attgattgctagatgactcactctccagttcaagaccacacttctactaacacacattt 1025
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 58 ATTTATTATGCTTGAATTTCTCGAGAGTACTGATTTCTGCCTCCATCCCACTCT 117
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 1026 ggaactctaacccagctacttcccattgaagcactactttgtaagctgactctctcaaga 1085
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 118 ATACACCTCCCTCCACTTCCAAATTAAGTGGTAATCTAGCCACTTACTGAATCTTCC 177
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 1086 caacactagggtctatttcttcgcgtttgggttggtacacgggtactaagccattgtc 1145
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 178 AGATAGAGAGCTCTTAAGATGGTCTCCACCCCGCTTTTACACACCATATCAGAAAAAGTTCG 237
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 1146 tactactctgttgtaattattgaagaaactgacgggttactctgcttcttcttgactgttcc 1205
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 238 TTATATTACACGAAGCTCTCGAAATGAGGATGATGAAGTTCAAGATGTACTTTTAC 297
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 1206 attcgcctctagagcttactggttgaaatgatgcaatgtgaagctgaaaggaaccattggt 1265
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 298 CTGCGTACCTTTTGTATCACTTGGCTAAGAAACATAGTCAGACGTTTGAAAAACAAAGTA 357
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 1266 tagagtttgggttaacgacagattt 1290
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 358 AAAGGTATTCTTTAGTAAAGATT 382
  RESULT 12
  AQ140737/c
  LOCUS
  DEFINITION
    AQ140737 545 bp DNA GSS 24-SEP-1998
    HS_3124_A2_D10_T7 CIT Approved Human Genomic Sperm Library D Homo
    sapiens genomic clone Plate-3124 Col-20 Row-G, DNA sequence.
  ACCESSION
    AQ140737
  VERSION
    AQ140737.1 GI:3531390
  KEYWORDS
    GSS.
  SOURCE
    human.
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 545)
    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
    Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
    Hood,L.
    Sequence-tagged connectors: A sequence approach to mapping and
    scanning the human genome
    Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
    99380589
    Contact: Mahairas GG, Wallace JC, Hood L
    High Throughput Sequencing Center
    University of Washington
    401 Queen Anne Avenue North, Seattle, WA 98109, USA
    Tel: (206) 616-3618
    Fax: (206) 616-3887
    Email: jwallace@u.washington.edu
    Sequence Tagged Connector
    Plate: 3124 row: G Column: 20
    Class: BAC ends
    High quality sequence stop: 545.
    Location/Qualifiers
      1..545
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="plate-3124 Col-20 Row-G"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /sex="male"
      /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
      E-Coli DH10B"
  BASE COUNT 185 a 90 c 90 g 178 t 2 others
  ORIGIN

  Query Match 3.0%; Score 41.8; DB 224; Length 545;
  Best Local Similarity 54.0%; Pred. No. 0.36;
  Matches 129; Conservative 0; Mismatches 107; Indels 3; Gaps 2;

  QY 284 actctgcttgattgaagctattcaaaagacgctactcttcaagggtgaagctctt 343
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 474 AATCTGATTGGTTTTTAATAGTAAAGTACACTGGTGGAGACATGAGTACGATCTT 415
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 344 tcttgaagacttacaactacactttgggtgctgacgactgactccat-tcggtaacaa 402
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 414 TATTCATATATATCTCTTAATATGTTATGACTATTATATCATATGACAGCTCAG 355
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 403 caaatggttaactctgggtatttaagttcacagaagacgcttggctgagaagatt 462
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 354 TACAAGGTTAAGTAAACCATTTATTTCTATCTGTGAAACAAG--AATAACTAGATATTTA 297
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY 463 gttccattcattagagcttctggtctgacagagttattgctctgctgaaagttcat 521
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db 296 AAGCTATACATTTAGGCTTTCTGTTCTTCAAGGGTTATATCTTCTTCTAGGCTTCAT 238
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  RESULT 13
  AV129427
  LOCUS
  DEFINITION
    AV129427 231 bp mRNA EST 01-JUL-1999
    AV129427 Mus musculus c57BL/6J 11-day embryo Mus musculus cdNA
    clone 2700068B02, mRNA sequence.

```


Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C.,
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, I., Sugahara
 Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N.,
 Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE
 JOURNAL
 COMMENT

RIKEN Mouse ESTs
 Unpublished (1999)
 Contact: Chie Owa
 Genome Science Laboratory
 RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-9145
 Fax: 81-298-36-9098

Email: genom-res@rtc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length cDNA
 (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.

FEATURES

Location/Qualifiers

1..236
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1810032D19"
 /clone_lib="Mus musculus pancreas C57BL/6J adult"
 /sex="male"
 /tissue_type="pancreas"
 /dev_stage="adult"

BASE COUNT 56 a 47 C 69 g 64 t

ORIGIN

Query Match 2.9%; Score 41.2; DB 109; Length 236;
 Best Local Similarity 52.3%; Pred. NO. 0.42;
 Matches 91; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 1154 ctgttgaaatctattgaagaactgacggttactctgtcttcttgactgttcattcgctg 1213
 DB 45 CTGGTGGGCCCGCTCCAGCAGCTGGAGCTGGCGCTGCTGGTACTGCTTCATCCACTG 104
 QY 1214 ctgagacttacgttgaaatgatgaatgtgaagctgaagaagaccattggttagagttt 1273
 DB 105 CCGCGCCCCAGCTGAGGAGAGAAAGTGGAGGCAAGAGAGAGATTCGAGGATTTTG 164
 QY 1274 tggtaacgacagaggtgttccattgcacggttgtggtgttgacaagttgggta 1327
 DB 165 AAGATGACATGGGCTTCGTTTTTTTGTGACTAACTGTTTTTTGTTAAGTTGGGTA 218

Search completed: October 26, 2001, 18:33:06
 Job time: 6796 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 17:50:32 ; Search time 210.38 Seconds
(without alignments)
4190.388 Million cell updates/sec

Title: US-09-488-265-30

Perfect score: 1404

Sequence: 1 atggcggttcgttcgtct.....gggaagaatgttcgtctaa 1404

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

```

1: /SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1404	100.0	1404	20	AAZ31521
2	1404	100.0	1404	21	AAZ3234
3	1404	100.0	1404	21	AAZ59716
4	1388	98.9	1404	21	AAZ73293
5	1388	98.9	1426	21	AAZ73232
6	1388	98.9	1426	21	AAZ59642
7	1384.8	98.6	1404	21	AAZ3292
8	1327.2	94.5	1404	20	AAZ31520
9	1327.2	94.5	1404	21	AAZ73233
10	1327.2	94.5	1404	21	AAZ59715
11	1322.4	94.2	1404	21	AAZ73291

12	1319.2	94.0	1404	21	AAZ73290	Consensus phytase
13	1317.6	93.8	1426	20	AAZ27423	Ascomycete Consens
14	1317.6	93.8	1426	20	AAZ23022	Fungal phytase gen
15	1317.6	93.8	1426	21	AAZ73231	Consensus phytase
16	1317.6	93.8	1426	21	AAZ59637	DNA encoding phyt
17	1258.4	89.6	1426	20	AAZ31523	Consensus phytase
18	1258.4	89.6	1426	21	AAZ73236	Consensus phytase
19	1258.4	89.6	1426	21	AAZ59738	DNA encoding phyt
20	653.4	46.5	1404	20	AAZ31522	A. fumigatus phyt
21	653.4	46.5	1404	21	AAZ73235	Aspergillus fumiga
22	653.4	46.5	1404	21	AAZ59717	DNA encoding a mut
23	615.8	43.9	1404	18	AAZ65136	Aspergillus ficum
24	614.2	43.7	1404	12	AAQ11175	Chromosomal phytas
25	614.2	43.7	1404	20	AAZ27421	A. ficum phytase
26	611	43.5	1404	12	AAQ13878	Phytase gene, Asp
27	610.2	43.5	2363	15	AAQ58126	Phytase gene, Asp
28	610.2	43.5	2379	15	AAQ56944	A. niger phytase g
29	610.2	43.5	6756	12	AAQ11174	Sequence, from ove
30	610.2	43.5	6756	18	AAZ65137	Aspergillus ficum
31	610	43.4	1571	19	AAZ03144	Aspergillus ficum
32	610	43.4	1571	20	AAZ27422	A. fumigatus phyt
33	583.6	41.6	1515	18	AAZ96709	Aspergillus niger
34	576.8	41.1	1931	19	AAZ03142	Aspergillus nidula
35	576.8	41.1	1931	20	AAZ27413	A. nidulans phytas
36	558	39.7	1912	17	AAZ03743	Phytase chimeric g
37	539.4	38.4	1845	19	AAZ03143	Talaromyces thermo
38	537.8	38.3	1845	20	AAZ27416	T. thermophilus ph
39	528.6	37.6	2327	17	AAZ03736	Phytase gene, Asp
40	528.6	37.6	2327	20	AAZ27414	A. terreus phytase
41	527	37.5	1567	19	AAZ03145	Aspergillus terreu
42	511.2	36.4	1922	22	AAZ77028	P. hordei phytase D
43	510.6	36.4	1464	22	AAZ77030	P. hordei phytase c
44	510.6	36.4	1584	22	AAZ77029	Part of P. hordei p
45	414.2	29.5	4898	22	AAZ77031	P. hordei phytase c

ALIGNMENTS

RESULT 1

AAZ31521
ID AAZ31521 standard; DNA; 1404 BP.

AC AAZ31521;

XX 06-JAN-2000 (first entry)

XX Consensus phytase-10-thermo(3)-Q50T-K91A coding sequence.

XX Phytase; animal feed preparation; thermostable phytase; transgenic plant;
XX consensus sequence; ds.

OS Synthetic.

XX WO9948380-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-DK00154.

XX 23-MAR-1998; 98DK-0000407.

XX 19-JUN-1998; 98DK-0000806.

XX 18-SEP-1998; 98DK-0001176.

XX 22-JAN-1999; 99DK-0000091.

XX 22-JAN-1999; 99DK-0000093.

XX (NOVO) NOVO-NORDISK AS.

XX Petersen S;

XX WPI; 1999-591030/50.

XX P-PSDB; AAY43170.

PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 10; 71pp; English.

PS This sequence encodes the consensus phytase-10-thermo(3)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;

Query Match 100.0%; Score 1404; DB 20; Length 1404;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcggttgcgtgctactgtccattgccaacttgggttcgggtccacatccggtacc 60
 PS |||||
 CC |||||
 DB 1 atggcggttgcgtgctactgtccattgccaacttgggttcgggtccacatccggtacc 60
 QY 61 gcttgggttcctcggtgaactctcaactctgtgacactgttgacggttggttaccatgt 120
 PS |||||
 CC |||||
 DB 61 gcttgggttcctcggtgaactctcaactctgtgacactgttgacggttggttaccatgt 120
 QY 121 ttcccagaaattctcaacttgggtggtacatactctcattctctcttcttggcgtagcaa 180
 PS |||||
 CC |||||
 DB 121 ttcccagaaattctcaacttgggtggtacatactctcattctctcttcttggcgtagcaa 180
 QY 181 tctgtatttccagaggttccaaaggtgtgagattacttctcggttcaagtttgtct 240
 PS |||||
 CC |||||
 DB 181 tctgtatttccagaggttccaaaggtgtgagattacttctcggttcaagtttgtct 240
 QY 241 agacacggtgctagataccactctctcgtctgaagggtactctgctttgattgaa 300
 PS |||||
 CC |||||
 DB 241 agacacggtgctagataccactctctcgtctgaagggtactctgctttgattgaa 300
 QY 301 gctattcaaaagacgctactcttccaaaggtgaagtagtcgttttttgaagacttacaac 360
 PS |||||
 CC |||||
 DB 301 gctattcaaaagacgctactcttccaaaggtgaagtagtcgttttttgaagacttacaac 360
 QY 361 tacacttgggtgctgacgacttgcctcattcgtgtaacaaatggttaactctggt 420
 PS |||||
 CC |||||
 DB 361 tacacttgggtgctgacgacttgcctcattcgtgtaacaaatggttaactctggt 420
 QY 421 attaatgtctacagaagatcacaggcttggctagaaagattgttccattcattagact 480
 PS |||||
 CC |||||
 DB 421 attaatgtctacagaagatcacaggcttggctagaaagattgttccattcattagact 480
 QY 481 tctggttctgacagagttattctctctgctgaaaagttcattgaaggtttccactctgct 540
 PS |||||
 CC |||||
 DB 481 tctggttctgacagagttattctctctgctgaaaagttcattgaaggtttccactctgct 540
 QY 541 aagttggtgacccaggtgctaacccacacacacacacacacacacacacacacacacac 600
 PS |||||
 CC |||||
 DB 541 aagttggtgacccaggtgctaacccacacacacacacacacacacacacacacacacac 600
 QY 601 ccagaaggtgctggttac 660
 PS |||||
 CC |||||
 DB 601 ccagaaggtgctggttac 660
 QY 661 tctgaattgggtgacagcttgaagctaacctcactgctgttttgcgtccacacattaga 720
 PS |||||
 CC |||||
 DB 661 tctgaattgggtgacagcttgaagctaacctcactgctgttttgcgtccacacattaga 720
 QY 721 gctagattggaagctcacttgcaggtgttaacttgactgacgacagagcttgttaacttg 780
 PS |||||
 CC |||||
 DB 721 gctagattggaagctcacttgcaggtgttaacttgactgacgacagagcttgttaacttg 780

QY 781 atggacatgttccattgacacactgttgcctagaacttctgcagctactcaattgtctcca 840
 PS |||||
 CC |||||
 DB 781 atggacatgttccattgacacactgttgcctagaacttctgcagctactcaattgtctcca 840
 QY 841 tctgtgacttctcaactcacgacgaatggattcaatacagactacttgcattcttgggt 900
 PS |||||
 CC |||||
 DB 841 tctgtgacttctcaactcacgacgaatggattcaatacagactacttgcattcttgggt 900
 QY 901 aagtactacggttaagctgctgttaacccattgggtccagctcagaggttgggttctgt 960
 PS |||||
 CC |||||
 DB 901 aagtactacggttaagctgctgttaacccattgggtccagctcagaggttgggttctgt 960
 QY 961 aacgaattgattgctagattgactcactcctcagttccaagaccacacacttctactaacac 1020
 PS |||||
 CC |||||
 DB 961 aacgaattgattgctagattgactcactcctcagttccaagaccacacacttctactaacac 1020
 QY 1021 acttggactctaaacccagctacttccattggaacgctacttctgtaagctgacttctct 1080
 PS |||||
 CC |||||
 DB 1021 acttggactctaaacccagctacttccattggaacgctacttctgtaagctgacttctct 1080
 QY 1081 cagcaacacactatggttcttctatttcttcgttgggtttgtacaacggtacttaagcca 1140
 PS |||||
 CC |||||
 DB 1081 cagcaacacactatggttcttctatttcttcgttgggtttgtacaacggtacttaagcca 1140
 QY 1141 ttgtctactacttctgttgaatctattgaagaactgaacggttactctctcttctggact 1200
 PS |||||
 CC |||||
 DB 1141 ttgtctactacttctgttgaatctattgaagaactgaacggttactctctcttctggact 1200
 QY 1201 gtccattgctgctagagcttgaatgataatgataatgataatgataatgataatgataatg 1260
 PS |||||
 CC |||||
 DB 1201 gtccattgctgctagagcttgaatgataatgataatgataatgataatgataatgataatg 1260
 QY 1261 ttggttagagtttgggttaacacagagttgttccattgcacggttgggttgggtgacaaag 1320
 PS |||||
 CC |||||
 DB 1261 ttggttagagtttgggttaacacagagttgttccattgcacggttgggttgggtgacaaag 1320
 QY 1321 ttgggttagagttgaagagacagacttgcgttgaaggtttgttcttcgctagatctggtgt 1380
 PS |||||
 CC |||||
 DB 1321 ttgggttagagttgaagagacagacttgcgttgaaggtttgttcttcgctagatctggtgt 1380
 QY 1381 aactgggaagaatgttccgttaa 1404
 PS |||||
 CC |||||
 DB 1381 aactgggaagaatgttccgttaa 1404

RESULT 2

AAA73234

ID AAA73234 standard; DNA; 1404 BP.

AC AAA73234;

XX 05-DEC-2000 (first entry)

DT Consensus phytase 10 thermo 3 q50t, k91a polynucleotide SEQ ID NO:30.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;

XX temperature stability; pH profile; temperature profile; reaction rate;

KW specific activity; substrate specificity; substrate cleavage pattern;

KW substrate binding; position specificity; phytate degradation rate;

KW food; feed; phytate; manure; ds.

XX Synthetic.

OS WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

PA

XX

PI Lehmann M;

XX WPI; 2000-491161/43.

DR P-PSDB; AAB20527.

XX Novel phytases with improved properties such as temperature stability,

PT pH stability and substrate specificity, for use in pharmaceuticals and
PT compound foods and feeds -

XX

PS Claim 10; Fig 8a-c; 24Opp; English.

XX

CC The present invention describes improved phytases, preferably with
CC increased thermostability, and methods for producing them. The methods
CC can be used for producing phytases with improved properties e.g. profile,
CC temperature stability, pH stability, pH profile, temperature profile,
CC specific activity, substrate specificity, substrate cleavage pattern,
CC substrate binding, position specificity, the velocity and level of
CC release of phosphate from corn, reaction rate, phytate degradation rate,
CC and end level of released phosphate. The phytases can be used to produce
CC pharmaceutical compositions or compound food or feeds. The feed can be
CC used to reduce levels of phytate in animal manure, by converting it
CC into lower inositol phosphates and/or inositol and inorganic phosphate.
CC The present sequence encodes a phytase sequence from the present
CC invention.

XX

SQ Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;

Query Match 100.0%; Score 1404; DB 21; Length 1404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgggggtgttcgtgctactgtccattgccaccctgttcggtccacatccggtacc 60

Db 1 atgggggtgttcgtgctactgtccattgccaccctgttcggtccacatccggtacc 60

Qy 61 gccitgggtctcgtggaactctcatctctgtgacactgttgacggtgttaccatgt 120

Db 61 gccitgggtctcgtggaactctcatctctgtgacactgttgacggtgttaccatgt 120

Qy 121 ttccacagaattctcactgtggtgacactctcatctctgtgacactgttgacggtgttaccatgt 180

Db 121 ttccacagaattctcactgtggtgacactctcatctctgtgacactgttgacggtgttaccatgt 180

Qy 181 tctgctattctccagacgttccaaagggtgttagagttacttctgctcaagtttctgt 240

Db 181 tctgctattctccagacgttccaaagggtgttagagttacttctgctcaagtttctgt 240

Qy 241 agacacgtgttagataccacactctctgctctaaagggtgttagagttacttctgctcaagtttctgt 300

Db 241 agacacgtgttagataccacactctctgctctaaagggtgttagagttacttctgctcaagtttctgt 300

Qy 301 gctattccaaagaacgtactgttccaaagggtgttagagttacttctgctcaagtttctgt 360

Db 301 gctattccaaagaacgtactgttccaaagggtgttagagttacttctgctcaagtttctgt 360

Qy 361 tacaatttggtgctgacgactgaactcattcgtggaacacaaatggttaactctgt 420

Db 361 tacaatttggtgctgacgactgaactcattcgtggaacacaaatggttaactctgt 420

Qy 421 attaagttctacagaagatcacaaggttggctagaaagattgttcattcattagagct 480

Db 421 attaagttctacagaagatcacaaggttggctagaaagattgttcattcattagagct 480

Qy 481 tctggtctgacagagtattgtctgtgctgaaagggttcaattgaaagtttccaaatctgt 540

Db 481 tctggtctgacagagtattgtctgtgctgaaagggttcaattgaaagtttccaaatctgt 540

Qy 541 aagttgctgacacaggtgtcaccacacacacagcttccagtttatttaacgttatt 600

Db 541 aagttgctgacacaggtgtcaccacacacacagcttccagtttatttaacgttatt 600

Qy 601 ccagaagtgctggtttacaaacactttggacacaggtttgtgtactgcttttcgaaga 660

Db 601 ccagaagtgctggtttacaaacactttggacacaggtttgtgtactgcttttcgaaga 660

Qy 661 tctgaattgggtgacacagcttgaagctaaactcactgctgttttcgctccaccaataga 720

Db 661 tctgaattgggtgacacagcttgaagctaaactcactgctgttttcgctccaccaataga 720

Qy 721 gctagattggaagctcacttgcaggtgttaacttgaactgaacgaagcgttgttaactgt 780

Db 721 gctagattggaagctcacttgcaggtgttaacttgaactgaacgaagcgttgttaactgt 780

Qy 781 atggacatgtgtcattcgacactgtgtgttagaacttctgaactcactcaattgtctcca 840

Db 781 atggacatgtgtcattcgacactgtgtgttagaacttctgaactcactcaattgtctcca 840

Qy 841 tctgtgactgttctactcagcagcaatggtatcattcaatacgaactcacttgcatttgggt 900

Db 841 tctgtgactgttctactcagcagcaatggtatcattcaatacgaactcacttgcatttgggt 900

Qy 901 aagtaactcaggttcaaggtgtgtgttaacccattgggtccagctcaaggttgtgttctgt 960

Db 901 aagtaactcaggttcaaggtgtgtgttaacccattgggtccagctcaaggttgtgttctgt 960

Qy 961 aacgaattgattgctagattgactcactctccagttccaaagacacacttctactaacca 1020

Db 961 aacgaattgattgctagattgactcactctccagttccaaagacacacttctactaacca 1020

Qy 1021 actttggactcctaaacacagctacttcccatggaacgctacttctgaactgaactctct 1080

Db 1021 actttggactcctaaacacagctacttcccatggaacgctacttctgaactgaactctct 1080

Qy 1081 cagcaacacactatggtttctcttctgtgttttgggtttgtacaacggttacttctgtgact 1140

Db 1081 cagcaacacactatggtttctcttctgtgttttgggtttgtacaacggttacttctgtgact 1140

Qy 1141 ttgtctactactctgtgaaactctattgaacttctcttctgtgttttgggtttgtacaacggttacttctgtgact 1200

Db 1141 ttgtctactactctgtgaaactctattgaacttctcttctgtgttttgggtttgtacaacggttacttctgtgact 1200

Qy 1201 gttccattcgtctgtagagcttacttgaactgtgaaatgtgaagctgaaagaaaccca 1260

Db 1201 gttccattcgtctgtagagcttacttgaactgtgaaatgtgaagctgaaagaaaccca 1260

Qy 1261 ttgggttagagtttgggttaacgacagaggtgttccattgcaaggttgggtgttgacaag 1320

Db 1261 ttgggttagagtttgggttaacgacagaggtgttccattgcaaggttgggtgttgacaag 1320

Qy 1321 ttgggttagagtttgggttaacgacagacttctgaaaggttcttctgtagatctgtggt 1380

Db 1321 ttgggttagagtttgggttaacgacagacttctgaaaggttcttctgtagatctgtggt 1380

Qy 1381 aactgggaagaatgttctgcttaa 1404

Db 1381 aactgggaagaatgttctgcttaa 1404

RESULT 3

AAZ59716

ID AAZ59716 standard; DNA; 1404 BP.

XX AC AAZ59716;

XX AC AAZ59716;

XX 19-APR-2000 (first entry)

XX DNA encoding a mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;

XX thermostable; animal feed; monogastric animal; phytate phosphorus;

XX phosphate availability; consensus; mutant; ds.

XX Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

QY	1	atgggcgtgttcgctgctgactctccattgccacattgttcggttcacatccgcgtacc	60
Db	1	atgggcgtgttcgctgctgactctccattgccacattgttcggttcacatccgcgtacc <td>60</td>	60
QY	61	gccttgggtctcgtgtgaactctcactctctgtgacactgttgacactgttgacactgt	120
Db	61	gccttgggtctcgtgtgaactctcactctctgtgacactgttgacactgttgacactgt	120
QY	121	ttccagaaattctcactgtggtggtacactctcactctcactctctctctctctctctct	180
Db	121	ttccagaaattctcactgtggtggtacactctcactctcactctctctctctctctctct	180
QY	181	tctgctatttccagacgttccaaaggtttagagttacttctgtaagttttgtctctct	240
Db	181	tctgctatttccagacgttccaaaggtttagagttacttctgtaagttttgtctctct	240
QY	241	agacacgtgtgatacccaacttctctgctcctcgtctcgtctcgtctcgtctcgtctcgt	300
Db	241	agacacgtgtgatacccaacttctctgctcctcgtctcgtctcgtctcgtctcgtctcgt	300
QY	301	gctattcaaaagaacgtactgttccaaaggtttagagttacttctgtaagttttgtctct	360
Db	301	gctattcaaaagaacgtactgttccaaaggtttagagttacttctgtaagttttgtctct	360
QY	361	tacactttgggtgtgacgttgaactccattcgtgtaacaaacaaatggttaactctggt	420
Db	361	tacactttgggtgtgacgttgaactccattcgtgtaacaaacaaatggttaactctggt	420
QY	421	attaagttctacagaagaacgttgcctagaagattgttccattcatttagagct	480
Db	421	attaagttctacagaagaacgttgcctagaagattgttccattcatttagagct	480
QY	481	tctggttctgacagagttattgcttctcgtgaaaggttccattgaaaggtttccaatctgt	540
Db	481	tctggttctgacagagttattgcttctcgtgaaaggttccattgaaaggtttccaatctgt	540
QY	541	asgttgggtgacccgggtgctaacccacacaaagcttctccagttattacagttattatt	600
Db	541	asgttgggtgacccgggtgctaacccacacaaagcttctccagttattacagttattatt	600
QY	601	ccagaaggtgctggtttacacacacactttggaccacggtttgtgactgctttcgaagaa	660
Db	601	ccagaaggtgctggtttacacacacactttggaccacggtttgtgactgctttcgaagaa	660
QY	661	tctgaattgggtgacgacgttgaaagcttaactcactcgtgttttcgctccacaaattaga	720
Db	661	tctgaattgggtgacgacgttgaaagcttaactcactcgtgttttcgctccacaaattaga	720
QY	721	gctagattggaagctcactccaggtgttaactgactgacgacgaagcgttgttaacttg	780
Db	721	gctagattggaagctcactccaggtgttaactgactgacgacgaagcgttgttaacttg	780
QY	781	atggacatgtgccattcgcacactgttgcagaaacttgcagcactcactcaattgtctcca	840
Db	781	atggacatgtgccattcgcacactgttgcagaaacttgcagcactcactcaattgtctcca	840
QY	841	ttctgtgacttgcactcagcgaatggttcaatacagcactcactgcaactcttgggt	900
Db	841	ttctgtgacttgcactcagcgaatggttcaatacagcactcactgcaactcttgggt	900
QY	901	aagtactcaggttacggtgctgttaacccattgggtccagctcaaggtgtgtgttcggt	960
Db	901	aagtactcaggttacggtgctgttaacccattgggtccagctcaaggtgtgtgttcggt	960
QY	961	aacgaattgattgctagattgactcactcctccagttcgaagaccacactctcactaacac	1020
Db	961	aacgaattgattgctagattgactcactcctccagttcgaagaccacactctcactaacac	1020
QY	1021	actttggacttaacccagctactttccattgaacgctactttgttacgctgactctct	1080
Db	1021	actttggacttaacccagctactttccattgaacgctactttgttacgctgactctct	1080

Aspergillus niger var. avamori.

Aspergillus niger str. NRRL3135.

Aspergillus fumigatus ATCC13073.

Aspergillus fumigatus ATCC32722.

Aspergillus fumigatus ATCC58128.

Aspergillus fumigatus ATCC26906.

Aspergillus fumigatus ATCC32239.

Emericella nidulans.

Talaromyces thermophilus ATCC20186.

Myceliophthora thermophila.

Paxillus involutus NN005693.

Trametes pubescens NN9343.

Agrocybe pediades NN009289.

Peniophora lycii NN006113.

Thermomyces lanuginosa.

Synthetic.

Key Location/Qualifiers

CDS 1..1404

/*tag= a

/product= "Phytase-10-thermo[3-Q50T-K91A"]

EP969089-A1.

05-JAN-2000.

23-JUN-1999; 99EP-0111949.

29-JUN-1998; 98EP-0111960.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Brugger R, Lehmann M, Wyss M;

WPI; 2000-099429/09.

New stabilised enzyme formulation, useful for feed compositions for monogastric animals -

Example 5; Fig 19; 101pp; English.

The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate phosphohydrolase) and one or more stabilising agents including xylitol or ribitol; polyethylene glycols with a molecular weight of 600 to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic, glutaric and succinic acid; carboxymethylcellulose; and sodium alginate. The stabilised phytase formulation is used in a method for preparing a feed composition for monogastric animals (e.g., pigs, poultry) and provides a monogastric animal with its dietary requirements of phosphorus. Although a large amount of phosphate is present in animal feed in the form of phytate phosphorus, monogastric animals are unable to utilise this form of phosphate, resulting in the addition of extra phosphate to the feed of such animals. Phytase enhances the nutritional value of plant material without the need for adding additional phosphate to the feed. The level of phosphate pollution in the environment is reduced by adding phytase to animal feed, as the animal can make use of the inorganic phosphate liberated from phytate phosphorus using the enzyme. The phytase formulation of the invention has an improved thermostability and can therefore remain stable during long-term storage and can withstand feed processing methods such as extrusion, expansion and pelleting. The present sequence represents DNA encoding a mutant phytase-10 consensus sequence, phytase-10-thermo[31-Q50T-K91A], which has a temperature optimum and melting point 4 degrees Celsius higher than that of phytase-10 (AAV69566). Its specific activity with phytate as a substrate is also strongly increased.

Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;

Query Match 100.0%; Score 1404; DB 21; Length 1404;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1404: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1081	cacgacaacactatggtttctattttcttcggtttggtttgttacacggtactaagcca	1140
Db	1081	cacgacaacactatggtttctattttcttcggtttggtttgttacacggtactaagcca	1140
Qy	1141	ttgtctactactctgttgaactctattgaagaaactgaagttactctgcttttggact	1200
Db	1141	ttgtctactactctgttgaactctattgaagaaactgaagttactctgcttttggact	1200
Qy	1201	gttcacattcgcctagagcttaacttgaatgatgcgaatgtgaagctgaaaaggaaacca	1260
Db	1201	gttcacattcgcctagagcttaacttgaatgatgcgaatgtgaagctgaaaaggaaacca	1260
Qy	1261	ttggtttagagtttttggtttaacgacagagttgttcatttgcaagttgtgtgttgacaag	1320
Db	1261	ttggtttagagtttttggtttaacgacagagttgttcatttgcaagttgtgtgttgacaag	1320
Qy	1321	ttgggttagagtgtaagagagagagacttcgttaagagttgttttcgcctagactcgtgtgt	1380
Db	1321	ttgggttagagtgtaagagagagagacttcgttaagagttgttttcgcctagactcgtgtgt	1380
Qy	1381	aactgggaagaatgttttcgcgttaa	1404
Db	1381	aactgggaagaatgttttcgcgttaa	1404

RESULT

RESOLUT	4	
AAA73293		
ID	AAA73293	standard; DNA; 1404 BP.
XX		
AC	AAA73293;	
XX		
DT	05-DEC-2000	(first entry)
XX		
DE	Consensus	phytase 10 thermo 5 Q50T K91A polynucleotide SEQ ID NO:96.
XX		
KW	Phytase;	mutant; thermostability; mutation; mutagenesis; pH stability;
KW	temperature	stability; pH profile; temperature profile; reaction rate;
KW	specific	activity; substrate specificity; substrate cleavage pattern;
KW	substrate	binding; position specificity; phytate degradation rate;
KW	food;	feed; phytate; manure; ds.
XX		
OS	Synthetic.	
XX		
PN	WO200043503-A1.	
XX		
PD	27-JUL-2000.	
XX		
XX	21-JAN-2000;	2000WO-DK00025.
XX		
XX	22-JAN-1999;	99DK-0000092.
PR	21-SEP-1999;	99DK-0001340.
PR		
XX		
PA	(NOVO)	NOVO NORDISK AS.
XX		
PI	Lehmann M;	
XX		
DR	WPI;	2000-491161/43.
DR	P-PSDB;	AAB20534.
XX		
XX	Novel	phytases with improved properties such as temperature stability,
PT	pH	stability and substrate specificity, for use in pharmaceuticals and
PT	compound	foods and feeds -
XX		
PS	Disclosure;	Fig 25a-c; 240pp; English.
XX		
CC	The	present invention describes improved phytases, preferably with
CC	increased	thermostability, and methods for producing them. The methods
CC	can	be used for producing phytases with improved properties e.g.
CC	temperature	stability, pH stability, pH profile, temperature profile,
CC	specific	activity, substrate specificity, substrate cleavage pattern,
CC	substrate	binding, position specificity, the velocity and level of
CC	release	of phosphate from corn, reaction rate, phytate degradation rate,
CC	and	end level of released phosphate. The phytases can be used to produce

QY 1 atggcggtgttgctgctactgtccattgccaccttggcttcggttccacatccggtaacc 60
DB 12 atggcggtgttgctgctactgtccattgccaccttggcttcggttccacatccggtaacc 71
QY 61 gctctgggtccctggttaactccactctgtgacactgttgacgtgggttaccatgt 120
DB 72 gctctgggtccctggttaactccactctgtgacactgttgacgtgggttaccatgt 131
QY 121 ttccagaaattctccactgtgggttacatctctccattctctctcttcttggctgaagaa 180
DB 132 ttccagaaattctccactgtgggttacatctctccattctctcttcttggctgaagaa 191
QY 181 tctgtatttctccagacgtttccaaagggttgtagagttaactttcttggctgaagtttctct 240
DB 192 tctgtatttctccagacgtttccaaagggttgtagagttaactttcttggctgaagtttctct 251
QY 241 agacacggtgtgatacccaactctctctgctcgaaggtgactctgcttggattgaa 300
DB 252 agacacggtgtgatacccaactctctctgctcgaaggtgactctgcttggattgaa 311
QY 301 gctattcaaaagacgctactgtcttccaagggttaagtagcgtttcttgaagacttacac 360
DB 312 gctattcaaaagacgctactgtcttccaagggttaagtagcgtttcttgaagacttacac 371
QY 361 tacacttgggtgctgaagacttgactccattcgttgacacacaaatggttaactctggt 420
DB 372 tacacttgggtgctgaagacttgactccattcgttgacacacaaatggttaactctggt 431
QY 421 attaatctacagaagatcacaggcttggctgaagaagattgtccattcattagagct 480
DB 432 attaatctacagaagatcacaggcttggctgaagaagattgtccattcattagagct 491
QY 481 tctggttgcagacagttattgtctgtgaaaggcttcattgaagggtttccattctgct 540
DB 492 tctggttgcagacagttattgtctgtgaaaggcttcattgaagggtttccattctgct 551
QY 541 aagttggctgacacaggtgtac 600
DB 552 aagttggctgacacaggtgtac 611
QY 601 ccagaaaggtgctgtgtac 660
DB 612 ccagaaaggtgctgt 671
QY 661 tctgaattgggtgacgacgttgaagctaatctcactgtgtgtgtgtgtgtgtgtgtgtgtgt 720
DB 672 tctgaattgggtgacgacgttgaagctaatctcactgtgtgtgtgtgtgtgtgtgtgtgtgt 731
QY 721 gctagattggaagctcacttgcaggtgttaactgtgactgacgacgacgttgaacttg 780
DB 732 gctagattggaagctcacttgcaggtgttaactgtgactgacgacgacgttgaacttg 791
QY 781 atggacatgtgtcattcgcacactgtgtgtagaacttctgacgctactcaattgtctcca 840
DB 792 atggacatgtgtcattcgcacactgtgtgtagaacttctgacgctactcaattgtctcca 851
QY 841 tctgtgacttgttccactcagcagcaatggattccatagactacttgcaattttgggt 900
DB 852 tctgtgacttgttccactcagcagcaatggattccatagactacttgcaattttgggt 911
QY 901 aagttactacggttacggt 960
DB 912 aagttactacggttacggt 971
QY 961 aacgaattgtgtgtagattgactcactctccagttccagacacacacacttctactaacac 1020
DB 972 aacgaattgtgtgtagattgactcactctccagttccagacacacacacttctactaacac 1031
QY 1021 actttgactctaacccagctactttccattgaagcgtactttgtacgctgactctct 1080
DB 1032 actttgactctaacccagctactttccattgaagcgtactttgtacgctgactctct 1091
QY 1081 caccgaacacactatggtttctattttcttctgcttgggtttgtacacggttactaagcca 1140

DB 1092 cagcaacaactatggtttctattttctcgtcttgggtttgtacaacggtactaagcca 1151
QY 1141 ttgtctactactctgttgaatctattgaagaactgacggttactctgtcttctggact 1200
DB 1152 ttgtctactactctgttgaatctattgaagaactgacggttactctgtcttctggact 1211
QY 1201 gtccattcgtcctagacgttgaatggaatgacgaatggaagctgaaaggaacca 1260
DB 1212 gtccattcgtcctagacgttgaatggaatgacgaatggaagctgaaaggaacca 1271
QY 1261 ttggttagagattttgttaacgacagagttgttccattgcacggttgggtgtgacaag 1320
DB 1272 ttggttagagattttgttaacgacagagttgttccattgcacggttgggtgtgacaag 1331
QY 1321 ttggttagagattttgttaacgacagcttgcgttgaagggtttgtcttctcgtcagatcgtggt 1380
DB 1332 ttggttagagattttgttaacgacagcttgcgttgaagggtttgtcttctcgtcagatcgtggt 1391
QY 1381 aactgggaagaattttcgttaa 1404
DB 1392 aactgggaagaattttcgttaa 1415
RESULT 7
AAA73292
ID AAA73292 standard; DNA; 1404 BP.
XX AC AAA73292;
XX XX
DT 05-DEC-2000 (first entry)
XX
DE Consensus phytase 10 thermo 5 Q50T polynucleotide SEQ ID NO:94.
XX
KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure; ds.
XX
OS Synthetic.
XX
PN WO200043503-A1.
XX
PD 27-JUL-2000.
XX
PR 21-JAN-2000; 2000WO-DK000025.
XX
PR 22-JAN-1999; 99DK-00000092.
PR 21-SEP-1999; 99DK-0001340.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Lehmann M;
XX
DR WPI; 2000-491161/43.
DR P-PSDB; AAB20533.
XX
PT Novel phytases with improved properties such as temperature stability,
PT pH stability and substrate specificity, for use in pharmaceuticals and
PT compound foods and feeds -
XX
PS Disclosure; Fig 24a-c; 240pp; English.
XX
CC The present invention describes improved phytases, preferably with
CC increased thermostability, and methods for producing them. The methods
CC can be used for producing phytases with improved properties e.g.
CC temperature stability, pH stability, pH profile, temperature profile,
CC specific activity, substrate specificity, substrate cleavage pattern,
CC substrate binding, position specificity, the velocity and level of
CC release of phosphate from corn, reaction rate, phytate degradation rate,
CC and end level of released phosphate. The phytases can be used to produce
CC pharmaceutical compositions or compound food or feeds. The feed can be

XX This sequence encodes the consensus phytaase-1-thermo(8)-Q50T-K91A.
CC The invention relates to a process for preparing animal feed by
CC agglomerating feed ingredients with a thermostable phytase, which is
CC added before or during agglomeration. The thermostable phytase is useful
CC for expression in transgenic plants. These plants are useful in the
CC preparation of animal feed itself. The thermostable phytase allows animal
CC feed to be produced more efficiently, in addition to improved
CC phytase-expressing transgenic plants. These plants provide a feed
CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;
SQ

Query Match 94.5%; Score 1327.2; DB 20; Length 1404;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1356; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 atggcggtttcgtcgctactgccaattgcacattgttcggtccacatcgcgtacc 60
| | | | |
Db 1 atggcggtttcgtcgctactgccaattgcacattgttcggtccacatcgcgtacc 60
| | | | |

QY 61 gccttggttcctcgtggttaattctcaactctgtgacactgttacggtgttaccagt 120
| | | | |
Db 61 gccttggttcctcgtggttaattctcaactctgtgacactgttacggtgttaccagt 120
| | | | |

QY 121 ttcccagaattctcaacttgtgggttacatactctccattctctcttggctgcagaa 180
| | | | |
Db 121 ttcccagaattctcaacttgtgggttacatactctccattctctcttggcagagaa 180
| | | | |

QY 181 tctgtattttccacagcgttccaaaagggttgtagactattcttcgaagttttgtct 240
| | | | |
Db 181 tctgtattttccacagcgttccacaagcgtgtagattactttcgttcaagttttgtct 240
| | | | |

QY 241 agacacggtgctagataccccaaactttctcgtctaaaggcgtacctgtgttgattgaa 300
| | | | |
Db 241 agacacggtgctagataccccaaactttctcgtctaaaggcgtacctgtgttgattgaa 300
| | | | |

QY 301 gctattcaaaaagaacgcgtactgctttcaaagggttaagtacgctttcttgaagacttacaa 360
| | | | |
Db 301 gctattcaaaaagaacgcgtactgctttcaaagggttaagtacgctttcttgaagacttacaa 360
| | | | |

QY 361 taacatttgggtgctgcagacttgactccattcgttgtaaagcgtacctgtttaactcgtgt 420
| | | | |
Db 361 taacatttgggtgctgcagacttgactccattcgttgtaaagcgtacctgtttaactcgtgt 420
| | | | |

QY 421 attaaagtctcacagaagatacaaaaggcttggctgaaagattgttcattcatagagct 480
| | | | |
Db 421 attaaagtctcacagaagatacaaaaggcttggctgaaagattgttcattcatagagct 480
| | | | |

QY 481 tctggtttctgcagagattattgcttctcgtgtaaaagtttcattgaagtttccaatcgtct 540
| | | | |
Db 481 tctggtttctgcagagattattgcttctcgtgtaaaagtttcattgaagtttccaatcgtct 540
| | | | |

QY 541 aaagtgtgctgccacaggctctaaccaccaaaccaagcttctccagttattaacggttatatt 600
| | | | |
Db 541 aaagtgtgctgccacaggcttctcaaccaccaaaccaagcttctccagttattaacggttatatt 600
| | | | |

QY 601 ccagaagggtcgtggtttacaacacacttggacacaggtttgtgactgcttttcgaagaa 660
| | | | |
Db 601 ccagaaggatccggtttacaacacacttggacacaggttactgtactgttttcgaagaa 660
| | | | |

QY 661 tctgaattgggtgacagacggttgaaagcttaacttcaactcgtctttttgcgtccacaattaga 720
| | | | |
Db 661 tctgaattagggtgacagacggttgaaagcttaacttcaactcgtctttttgcgtccagctattaga 720
| | | | |

QY 721 gctagaattggaagctcgaactgtccacagggtttaacttgactgacgaagacggttgttaactgt 780
| | | | |
Db 721 gctagaattggaagctgaactgtccacagggtttaacttgactgacgaagacggttgttactgtg 780
| | | | |

QY 781 atggacatgtgtccattcgcacacttgttgcgtagaaacttctgacgtcactcaattgtctoca 840
| | | | |
Db 781 atggacatgtgtccattcgcacacttgtcgtagaaacttctgacgtcactcaattgtctoca 840
| | | | |

Qy	841	tctgtgactgtttcactcacgacgaatggattcaatacagctactactcgcgaactctttgggt	900
Db	841		
Qy	841	tctgtgtttgttcaactcacgacgaatggattcccaatacagctactctgcgaagcttgggt	900
Db	841		
Qy	901	aagtactacggtttacggtgctggttaaccatttgggttcagctcaaggtgttgggtttcgtt	960
Db	901		
Qy	901	aagtactacggtttacggtgctggttaaccatttgggttcagctcaaggtgttgggtttcgtt	960
Db	901		
Qy	961	aacgaattgattgtcgtagattgactcactctccagttccagaccacactcttactaaccac	1020
Db	961		
Qy	961	aacgaattgattgtcgtagattgactcactctccagttccagaccacactcttactaaccac	1020
Db	961		
Qy	1021	actttggactctaacccagctactttccoaattgaacgctactttgttaacgctgactctct	1080
Db	1021		
Qy	1021	actttggactctaacccagctactttccoaattgaacgctactttgttaacgctgactctct	1080
Db	1021		
Qy	1081	cacgacaacactatggtttctatattttcttcgctttgggtttgtacaacggttactaagcca	1140
Db	1081		
Qy	1081	cacgacaacactatgatactatattttcttcgctttgggtttgtacaacggttactaagcca	1140
Db	1081		
Qy	1141	ttgttactactctctgtgaactctattgaagaaactcagcgttactctgcttcttgact	1200
Db	1141		
Qy	1141	ttgttactactctgtgtgaactctattgaagaaactcagcgttactctgcttcttgact	1200
Db	1141		
Qy	1201	gttccattcgtcgtctagagcttaagttgaaatgatacgaatgtgaagctgaaaggaacca	1260
Db	1201		
Qy	1201	gttccattcgtcgtctagagcttaagttgaaatgatacgaatgtgaagctgaaaggaacca	1260
Db	1201		
Qy	1261	ttgggttagagttttggtttaacacacagagtgtgttccattgcagcgtttggtgttgacaag	1320
Db	1261		
Qy	1261	ttgggttagagttttggtttaacacacagagtgtgttccattgcagcgtttggtgttgacaag	1320
Db	1261		
Qy	1321	ttgggttagagttaagagagacgactcgttgtgaagttttgtcttcttcgctagatctgggtgt	1380
Db	1321		
Qy	1321	ttgggttagagttaagagagacgactcgttgtgaagttttgtcttcttcgctagatctgggtgt	1380
Db	1321		
Qy	1381	aactgggaagaatgtttcgccttaa	1404
Db	1381		
Qy	1381	aactgggcgtgaatgtttcgccttaa	1404
Db	1381		

RESULT	9	
AAA73233		
ID	AAA73233 standard; DNA; 1404 BP.	
XX		
XX		
AC	AAA73233;	
XX		
DT	05-DEC-2000 (first entry)	
XX		
DE	Consensus phytase 1 thermo 8 q50	
XX		
KW	Phytase; mutant; thermostability	
KW	temperature stability; pH profile	
KW	specific activity; substrate spec	
KW	substrate binding; position spec	
KW	food; feed; phytate; manure; ds.	
XX		
OS	Synthetic.	
XX		
PN	WO200043503-A1.	
XX		
PD	27-JUL-2000.	
XX		
PF	21-JAN-2000; 2000WO-DK00025.	
XX		
PR	22-JAN-1999; 99DK-0000092.	
PR	21-SEP-1999; 99DK-0001340.	
XX		
PA	(NOVO) NOVO NORDISK AS.	
XX		
PI	Lehmann M;	
XX		

OS Aspergillus fumigatus AFCC32722.
OS Aspergillus fumigatus AFCC58128.
OS Aspergillus fumigatus AFCC26906.
OS Aspergillus fumigatus AFCC32239.
OS Emericella nidulans.
OS Talaromyces thermophilus AFCC20186.
OS Myceliophthora thermophila.
OS Synthetic.
XX Key Location/Qualifiers
FH 1..1404
FT CDS /*tag= a
FT /product= "Phytase-1-thermo[8]-Q50T-K91A"
XX EP969089-A1.
XX 05-JAN-2000.
XX 23-JUN-1999; 99EP-0111949.
XX 29-JUN-1998; 98EP-0111960.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Brugger R, Lehmann M, Wyss M;
XX WPI: 2000-099429/09.
XX P-PSDB; AAY69568.
XX New stabilized enzyme formulation, useful for feed compositions for
XX monogastric animals.
XX Example 5; Fig 19; 101pp; English.
XX The invention relates to a novel stabilised dry or liquid enzyme
XX formulation, comprising phytase (myo-inositol hexakisphosphate
XX phosphohydrolase) and one or more stabilising agents including
XX xylitol or ribitol; polyethylene glycols with a molecular weight of 600
XX to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
XX glutamic and succinic acid; carboxymethylcellulose; and sodium alginate.
XX The stabilised phytase formulation is used in a method for preparing a
XX feed composition for monogastric animals (e.g., pigs, poultry) and
XX provides a monogastric animal with its dietary requirements of
XX phosphorus. Although a large amount of phosphate is present in animal
XX feed in the form of phytate phosphorus, monogastric animals are unable
XX to utilise this form of phosphate, resulting in the addition of extra
XX phosphate to the feed of such animals. Phytase enhances the nutritional
XX value of plant material without the need for adding additional phosphate
XX to the feed. The level of phosphate pollution in the environment is
XX reduced by adding phytase to animal feed, as the animal can make use of
XX the inorganic phosphate liberated from phytate phosphorus using the
XX enzyme. The phytase formulation of the invention has an improved
XX thermostability and can therefore remain stable during long-term storage
XX and can withstand feed processing methods such as extrusion, expansion
XX and pelleting. The present sequence represents DNA encoding a mutant
XX phytase-1 consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a
XX temperature optimum and melting point 7 degrees Celsius higher than that
XX of phytase-1 (AAY69558).
XX Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;
SQ

Query Match 94.5%; Score 1327.2; DB 21; Length 1404;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1356; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 1 atggcggtgttcgtcgctactgtccattgcccacattgttcggttcacacatccggtacc 60
DB 1 atggcggtgttcgtcgctactgtccattgcccacattgttcggttcacacatccggtacc 60
QY 61 gacctgggttcgtcggttaactctactctgtgacactgttgacggtgtggtaccatgt 120
DB 61 gacctgggttcgtcggttaactctactctgtgacactgttgacggtgtggtaccatgt 120

QY 121 ttccagaaatttctcacttgggtgtacatactctccattctctcttcttggctgacgaa 180
DB 121 ttccagaaatttctcacttgggtgtacatactctccattctctcttcttggcagacgaa 180
QY 181 tctgctatttctccagacgttccaaagggtgttagaggttacttcttcttcttcttctct 240
DB 181 tctgctatttctccagacgttccagacgttagaggttacttcttcttcttcttcttctct 240
QY 241 agacaggtgtctagatatacccaacttcttctgctctaaagcgtactctcttcttcttct 300
DB 241 agacaggtgtctagatatacccaacttcttctgctctaaagcgtactctcttcttcttct 300
QY 301 gctattcaaaaagacgtactctgttttcaagggttaagtacgcttcttcttcttcttcttct 360
DB 301 gctattcaaaaagacgtactctgttttcaagggttaagtacgcttcttcttcttcttcttct 360
QY 361 tacactttgggtgtcgtgacgacttgcatttcggttgaaacacaaatgttactctgtt 420
DB 361 tacactttgggtgtcgtgacgacttgcatttcggttgaaacacaaatgttactctgtt 420
QY 421 attaatctacagaagatacaaggcttctgctagaaagattgttccattctcattagagct 480
DB 421 attaatctacagaagatacaaggcttctgctagaaagattgttccattctcattagagct 480
QY 481 tctggttctgacagagtattgtcttctgctgaaaaagttcattgaaaggtttcccaatctgt 540
DB 481 tctggttctgacagagtattgtcttctgctgaaaaagttcattgaaaggtttcccaatctgt 540
QY 541 aagttggtgacccagggtgttaacccacacacacagcttctccagttatttaacgttattatt 600
DB 541 aagttggtgacccagggtgttccacacacacacagcttctccagttatttaacgttattatt 600
QY 601 ccagaaggtctggttac 660
DB 601 ccagaaggtctggttac 660
QY 661 tctgaattggtgacagcgttgaagcgttaacttcaacttcttcttcttcttcttcttcttcttct 720
DB 661 tctgaattggtgacagcgttgaagcgttgaacttcaacttcttcttcttcttcttcttcttcttct 720
QY 721 gctagattggaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgt 780
DB 721 gctagattggaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgttgaagcgt 780
QY 781 atgacatgtctccattcgac 840
DB 781 atgacatgtctccattcgac 840
QY 841 tctgtgactgttctcactcagcagcgaatggaatggaatggaatggaatggaatggaatggaatg 900
DB 841 tctgtgactgttctcactcagcagcgaatggaatggaatggaatggaatggaatggaatggaatg 900
QY 901 aagtactacggttactcgtggttgcac 960
DB 901 aagtactacggttactcgtggttgcac 960
QY 961 aacgaattgattgttagattgactcactctccagttcaagacacacacacacacacacacacacac 1020
DB 961 aacgaattgattgttagattgactcactctccagttcaagacacacacacacacacacacacacac 1020
QY 1021 accttgactatacccaagcgtacttccattgacacacacacacacacacacacacacacacacacac 1080
DB 1021 accttgactatacccaagcgtacttccattgacacacacacacacacacacacacacacacacacac 1080
QY 1081 caccgac 1140
DB 1081 caccgac 1140
QY 1141 ttgtctactactctctgttgaatctatttgaagaacacacacacacacacacacacacacacacacacac 1200
DB 1141 ttgtctactactctctgttgaatctatttgaagaacacacacacacacacacacacacacacacacacac 1200

QY 841 ttctgtgaacttgcactcaacgaagaatggattcaatacagactacttgcaactcttgggt 900
 Db 841 ttctgtgctttgttcaactcaacgaagaatggattcaatacagactacttgcaactcttgggt 900
 QY 901 aagtactacggttacggtgctgtaaccattgggtccagctcaagggtggttgcgtt 960
 Db 901 aagtactacggttacggtgctgtaaccattgggtccagctcaagggtggttgcgtt 960
 QY 961 aacgaattgattgctagattgactcactctccagttccagatccagaccacactcttactaacac 1020
 Db 961 aacgaattgattgctagattgactcactctccagttccagatccagaccacactcttactaacac 1020
 QY 1021 actttgactctaacccagctacttccattgaacgctactttgtacgctgactctct 1080
 Db 1021 actttgactctaacccagctacttccattgaacgctactttgtacgctgactctct 1080
 QY 1081 cagcaaacactatggttctctatttcttcgctttggtttgtacacaggtactaacgca 1140
 Db 1081 cagcaaacactatgattatctatttcttcgctttggtttgtacacaggtactaacgca 1140
 QY 1141 ttgtctactactctgttgaaatctattgaagaactgacggttactctgcttcttggact 1200
 Db 1141 ttgtctactactctgttgaaatctattgaagaactgacggttactctgcttcttggact 1200
 QY 1201 gtccattcgcgtgtagacttacgtttgaatgaaatgatgcaatgtgaagtgaaagaaacca 1260
 Db 1201 gtccattcgcgtgtagacttacgtttgaatgaaatgatgcaatgtgaagtgaaagaaacca 1260
 QY 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttgggttgacaaag 1320
 Db 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttgggttgacaaag 1320
 QY 1321 ttggttagatgtaagagacagactctgtgaaggtttgtcttcttcgtagatctgggt 1380
 Db 1321 ttggttagatgtaagagacagactctgtgaaggtttgtcttcttcgtagatctgggt 1380
 QY 1381 aactgggaagaatgttccgcttaa 1404
 Db 1381 aactgggctgaatgttccgcttaa 1404

RESULT 13
 ID AA227423
 AC AA227423 standard; cDNA; 1426 BP.
 AC AA227423;
 DT 07-DEC-1999 (first entry)
 DE Ascomycete consensus phytase coding sequence.
 KW Phytase; variant; enzyme; phosphorus liberation; phytase substrate;
 KW phytate level reduction; animal manure; food preparation;
 KW soy processing; inositol manufacture; ss.
 OS Synthetic.
 PN WO9949022-A1.
 PD 30-SEP-1999.
 PF 22-MAR-1999; 99WO-DK00153.
 PR 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 PA (NOVO) NOVO-NORDISK AS.
 XX Svendsen A;
 XX WPI; 1999-580444/49.

P-PSDB; AAY39906.

New variant phytase enzymes, used for liberating phosphorus from a phytase substrate, for reducing phytate levels in animal manure and in feed and food preparations -

Disclosure; Fig 9f-g; 141pp; English.

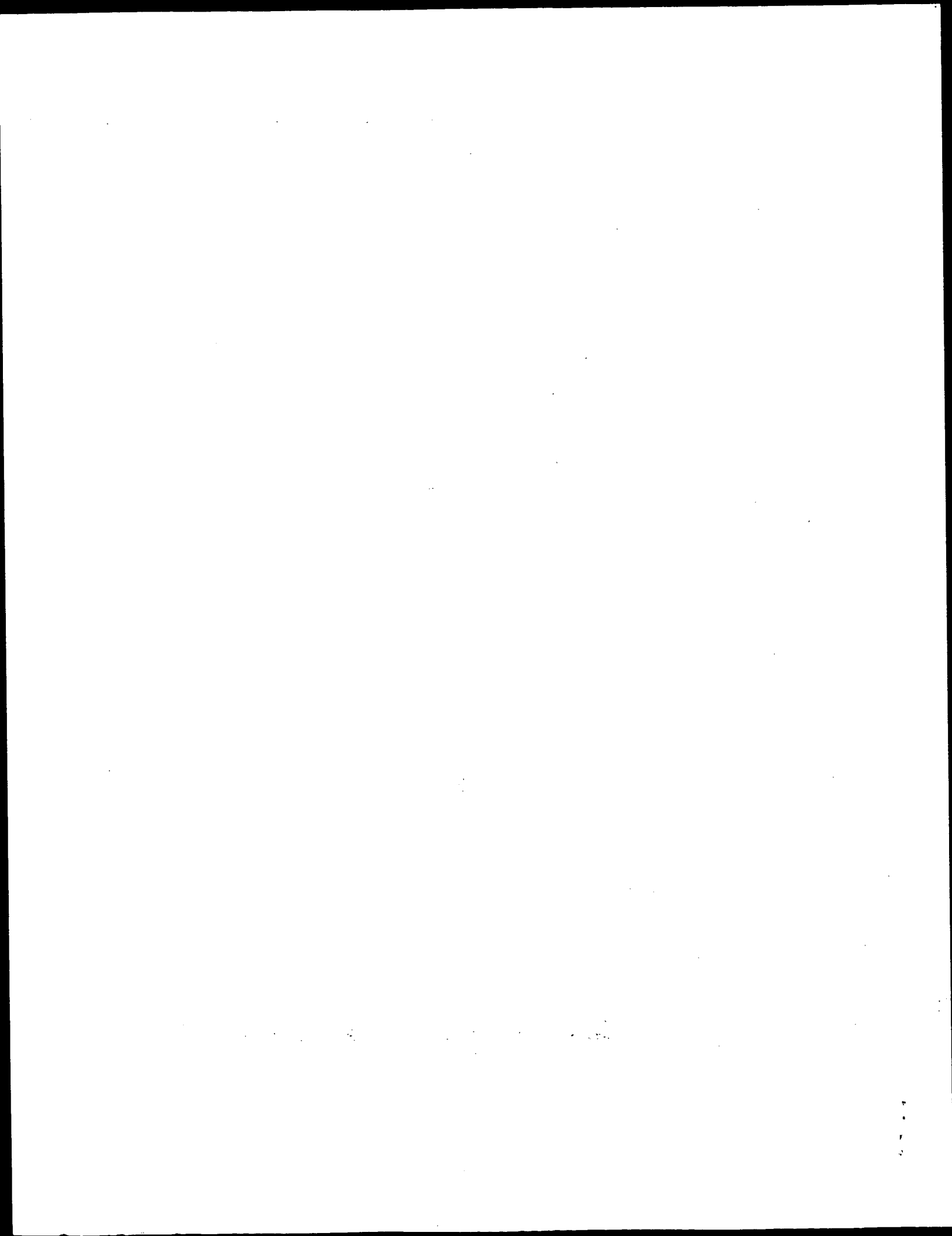
This sequence encodes the consensus Ascomycete phytase sequence. The invention relates to variant phytase enzymes with specific amino acid substitutions for improved properties. The phytase variants can be used for liberating phosphorus from a phytase substrate. They can be used for reducing phytate levels in animal manure. They can be used in feed or food preparations. The phytase DNA can also be used to produce transgenic plants which can be used in feeds or foods. The phytase variants can also be used in soy processing and in the manufacture of inositol or derivatives. The phytase variants can have altered activities such as pH stability, temperature stability, pH profile, temperature profile, specific activity (in particular in relation to pH and temperature), substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytase degradation rate and end level of released phosphate reached.

Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;

Query Match 93.8%; Score 1317.6; DB 20; Length 1426;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 1350; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 atggcggtgtctgctgctactgtccattgcacattgcttgcgttccacatccggtacc 60
 Db 12 atggcggtgtctgctgctactgtccattgcacattgcttgcgttccacatccggtacc 71
 QY 61 gccttgggtctcgtgtgtaactctcactctgtgacactgttgacggtgttaccactgt 120
 Db 72 gccttgggtctcgtgtgtaactctcactctgtgacactgttgacggtgttaccactgt 131
 QY 121 tcccgagaatttctcactgtggtggtggtggtggtggtggtggtggtggtggtggtggt 180
 Db 132 tcccgagaatttctcactgtggtggtggtggtggtggtggtggtggtggtggtggtggt 191
 QY 181 tctgtatttctccagacgttccaaagggtgttagagttacttcttcttcttcttcttct 240
 Db 192 tctgtatttctccagacgttccaaagggtgttagagttacttcttcttcttcttcttct 251
 QY 241 agacacggtgtgtagatatacccaacttcttctgctctaaagggttacttcttcttcttct 300
 Db 252 agacacggtgtgtagatatacccaacttcttctgctctaaagggttacttcttcttcttct 311
 QY 301 gctattcaaaagacgctactcttccaaagggtgttagagttacttcttcttcttcttcttct 360
 Db 312 gctattcaaaagacgctactcttccaaagggtgttagagttacttcttcttcttcttcttct 371
 QY 361 tacacttgggtgctgacgactgactcctcctcctcctcctcctcctcctcctcctcctcct 420
 Db 372 tacacttgggtgctgacgactgactcctcctcctcctcctcctcctcctcctcctcctcct 431
 QY 421 attagttctacagaagatacagaagcttcttggctagaagaagattgttccattcattagact 480
 Db 432 attagttctacagaagatacagaagcttcttggctagaagaagattgttccattcattagact 491
 QY 481 tctgttctgacagagttattgtcttctcgtctgaaggttcaattgaaggttcccaatctgct 540
 Db 492 tctgttctgacagagttattgtcttctcgtctgaaggttcaattgaaggttcccaatctgct 551
 QY 541 aagtgtggtgacctgactgacttaaccacacacacacacacacacacacacacacacacacac 600
 Db 552 aagtgtggtgacctgactgacttaaccacacacacacacacacacacacacacacacacacac 611
 QY 601 ccagaaggtgtgttatacaacaacacttggaccacggttctgttactgcttctcgaagaa 660
 Db 601 ccagaaggtgtgttatacaacaacacttggaccacggttctgttactgcttctcgaagaa 660

D	b	312	g	t	a	t	t	c	a	a	a	a	a	a	c	g	t	a	c	t	g	t	t	t	c	a	a	g	g	t	a	a	g	t	a	a	g	t	t	c	a	a	g	t	t	c	a	a	c	t	t	a	a	c		371
Q	y	361	t	a	c	a	t	t	i	g	g	t	g	t	g	c	a	g	a	c	t	g	a	c	c	a	t	c	c	a	t	c	o	g	g	t	a	a	c	a	a	a	a	a	t	t	a	a	c	t	t	g	t		420	
D	b	372	t	a	c	a	t	t	i	g	g	t	g	t	g	c	a	g	a	c	t	g	a	c	c	a	t	c	c	a	t	c	o	g	g	t	a	a	c	a	a	a	a	a	a	t	t	a	a	c	t	t	g	t		431
Q	y	421	a	t	t	a	a	g	t	t	c	a	a	a	a	a	t	a	c	a	a	g	t	t	i	g	g	t	g	t	a	a	a	a	g	a	a	t	t	t	c	a	t	t	c	a	t	t	a	a	g	t		480		
D	b	432	a	t	t	a	a	g	t	t	c	a	a	a	a	a	t	a	c	a	a	g	t	t	i	g	g	t	g	t	a	a	a	a	g	a	a	t	t	t	c	a	t	t	c	a	t	t	a	a	g	t		491		
Q	y	481	t	c	t	g	t	t	c	a	a	a	g	t	t	a	t	t	c	o	g	t	g	t	g	t	g	t	g	a	a	a	a	a	g	t	t	c	a	a	t	t	c	a	a	t	t	c	t		540					
D	b	492	t	c	t	g	t	t	c	a	a	a	g	t	t	a	t	t	c	o	g	t	g	t	g	t	g	t	g	a	a	a	a	a	g	t	t	c	a	a	t	t	c	a	a	t	t	c	t		551					
Q	y	541	a	a	g	t	t	g	t	g	a	c	c	a	a	g	t	g	t	a	a	c	c	a	c	a	c	a	a	c	a	a	a	a	g	t	t	c	o	a	g	t	t	c	a	a	c	t	t		600					
D	b	552	a	a	g	t	g	t	g	a	c	c	a	a	g	t	t	c	a	a	c	a	c	a	c	a	c	a	a	c	a	a	a	a	g	t	t	c	o	a	g	t	t	c	a	a	c	t	t		611					
Q	y	601	c	a	a	a	a	g	t	g	t	t	a	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a		660				
D	b	612	c	a	a	a	a	g	t	a	c	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a		671					
Q	y	661	t	c	t	g	a	a	t	t	i	g	g	t	g	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a		720				
D	b	672	t	c	t	g	a	a	t	t	i	g	g	t	g	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a		731				
Q	y	721	g	e	t	a	a	t	t	g	a	a	g	t	c	a	c	t	c	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a		780					
D	b	732	g	e	t	a	a	t	t	g	a	a	g	t	g	a	c	a	c	a	c</																																			



Db 312 gctattcaaaagacgctactgcttttcaagggttaagtaagctgttttgaagacttaacaac 371
Qy 361 tacactttgggtgctgacgactgactccattcggttgaaacacaaatgggttaactctggt 420
Db 372 tacactttgggtgctgacgactgactccattcggttgaaacacaaatgggttaactctggt 431
Qy 421 attaatgtctacagaagatacaaaagccttggctagaagaattgttccattcattagagct 480
Db 432 attaatgtctacagaagatacaaaagccttggctagaagaattgttccattcattagagct 491
Qy 481 tctggttctgacagagttattgcttctgctgaaagttcattgaaagtttcccaatctgct 540
Db 492 tctggttctgacagagttattgcttctgctgaaagttcattgaaagtttcccaatctgct 551
Qy 541 aagttgggtgacccaggtgtctaacccacacacaaagcttctccagttataaacggttattatt 600
Db 552 aagttgggtgacccaggtgtctaacccacacacaaagcttctccagttattgacggttattatt 611
Qy 601 ccagaaggtgctgtttac 660
Db 612 ccagaaggtgctgtttac 671
Qy 661 tctgaattgggtgacgacgcttgaagcttaacttcaacttcaacttctgcttccaccacattaga 720
Db 672 tctgaattgggtgacgacgcttgaagcttaacttcaacttcaacttctgcttccaccacattaga 731
Qy 721 gctagattggaagctcaacttccaggtgttcaacttgacgacgacgacgacgacgacgacgac 780
Db 732 gctagattggaagctcaacttccaggtgttcaacttgacgacgacgacgacgacgacgacgac 791
Qy 781 atggacatgttccattcgacacacttctgctagaacttctgacgctactcaacttctctccca 840
Db 792 atggacatgttccattcgacacacttctgctagaacttctgacgctactcaacttctctccca 851
Qy 841 ttctgactgttcaactcaacgacgaatggattcaactcaactcaactcaactcaactcaactcaact 900
Db 852 ttctgactgttcaactcaacgacgaatggattcaactcaactcaactcaactcaactcaactcaact 911
Qy 901 aagtaactcaggttgcgtgctgtgaac 960
Db 912 aagtaactcaggttgcgtgctgtgaac 971
Qy 961 aacgaattgattgctagattgactcaacttccagttcaacttcaacttcaacttcaacttcaacttca 1020
Db 972 aacgaattgattgctagattgactcaacttccagttcaacttcaacttcaacttcaacttcaacttca 1031
Qy 1021 actttggacttaacccagctacttcccaattgaacgctacttcttgaacttcttgaacttcttgaact 1080
Db 1032 actttggacttaacccagctacttcccaattgaacgctacttcttgaacttcttgaacttcttgaact 1091
Qy 1081 cagcaac 1140
Db 1092 cagcaac 1151
Qy 1141 ttgtctactacttctgttgaattctattgaagaacacggttacttctgttcttcttggact 1200
Db 1152 ttgtctactacttctgttgaattctattgaagaacacggttacttctgttcttcttggact 1211
Qy 1201 gttccattcgtctgtagaacttcaacttgaatgacgaatgacgaatgacgaatgacgaatgacgaat 1260
Db 1212 gttccattcgtctgtagaacttcaacttgaatgacgaatgacgaatgacgaatgacgaatgacgaat 1271
Qy 1261 ttggttagagttttggttaacacacagagttgttccattgacaggttgggtgttgaacag 1320
Db 1272 ttggttagagttttggttaacacacagagttgttccattgacaggttgggtgttgaacag 1331
Qy 1321 ttggttagatgttaagagacgacttctggttgaaggttcttcttcttcttcttcttcttcttcttct 1380
Db 1332 ttggttagatgttaagagacgacttctggttgaaggttcttcttcttcttcttcttcttcttcttct 1391
Qy 1381 aactgggaagaatgttctcgttaa 1404
Db 1392 aactgggtgaatgttctcgttaa 1415

RESULT 2

US-08-151-574-33
; Sequence 33, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Faridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Seilen
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 07/688,578
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; US-08-151-574-33

Query Match 43.7%; Score 614.2; DB 1; Length 1404;
Best Local Similarity 64.9%; Pred. No. 4.6e-173;
Matches 910; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

Qy 1 atggggcgttctgctgctactgctccattgacacactgttctgctgcttccacatccggtacc 60
Db 1 ATGGGCGTCTCTGCTGCTTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Qy 61 gcccttgggttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 120
Db 61 GCAGTCCCCCGCTCGAGAAATCAATCCAGTTCGATACGCTCGATCAGGGGTATCAATGC 120
Qy 121 ttcccagaaatttctcacttctggtgggttacatactctccattcttcttcttcttcttcttcttct 180
Db 121 TTCTCCGAGACTTCGCATCTTTGGGGTCAATAGCGACCGTTCTTCTCTCTCTGCGCAACGAA 180

Db 61 GCAGTCCCGCGCTCGAGAAATCAATCCAGTTCCGATACGGTCGATACGGGGTATCAATGC 120
QY 121 ttccagaaattctcaattgtgggtacatacttccattctctctcttctgttgctgaagaa 180
Db 121 TTCTCCGAGACTGTCATCTTTGGGGTCAATACGACCGTTCTTCTCTCTGCGCAAGAA 180
QY 181 tctgtatttctccagacgttccaaagggtttagagttacttctcgttcaagttttgtct 240
Db 181 TCGGTCACTCCCTGAGTGGCCGCGGATGCGAGTCACTTTTCGCTCAGTCTCTCTCC 240
QY 241 agacaggtgttagatataccaaattctctgtctgtctgaaggtactctgtttgattgaa 300
Db 241 CGTCATGGAGCGGTATCCGACCGACTCCCAAGGGCAAGAAATACTCGCGCTCTCATTTAG 300
QY 301 gctattcaaaagaaacgtactgtcttcaagggttaagtacgttcttgaagacttaaac 360
Db 301 GAGATCCAGCAGACCGCACCTTTGACGGAAATATGCTTCCGAGACATCAAC 360
QY 361 tacatttgggtgtgacgtacttgcctcattcgtgtaacaaatggttaactctggt 420
Db 361 TACAGCTTGGGTGACAGTACCTGACTCCCTTCGGAGACAGGAGTAGTCAACTCCGCG 420
QY 421 ataatgttcaagaagatacaaggcttgggtgtagaagattgttccattcattagagct 480
Db 421 ATCAAGTTCTACCAAGCGTACGATCGCTCAAGGAACATCGTTCCATTCATCCGATCC 480
QY 481 tctgttctgacagatttcttctgtctgtaaaagttcattgaagtttccaatctgct 540
Db 481 TCTGGCTCCAGCGCGGTGATCGCTCCGGCAAGAAATTCATCGAGGGCTCCAGAGCACC 540
QY 541 aagttgggtgacccaggtgtcttaaccacacacaaagttctcagttattaaagttattatt 600
Db 541 RAGCTGAAGGATCTCTGTCGCCAGCGGCCAATCGTCCGCCAAGATCGAGTGGTCAAT 600
QY 601 ccagaaggtgtggttacaacaaacttggacacaggttgtgtactgttcttccgaagaa 660
Db 601 TCCGAGGCGAGTCAATCAACAACACTTCGACCCAGGCACTGCATCTGTCTTCCGAAGAC 660
QY 661 tctgaattgggtgacaggttgaagtaacttaactcactgttctcgtcccaaataga 720
Db 661 AGCGAATTGGCGATACCGTGCAGACCAATTCACCGCCAGCTTCGTCCTCCATTCGT 720
QY 721 gctagattggaagctcaacttccaggtgttaactgtgactgacgaagacgttgttaacttg 780
Db 721 CAACGCTGGAGAGACCTGTCCGGTGTGACTCTCACAGACACAGAAGTACCTTACCTC 780
QY 781 atggacatgttccattcgaactgttctgtgtagaacttctgacgtactcgaattgtctcca 840
Db 781 ATGGACATGTGCTCTCGACACCATCTCCACACAGCAGCGTCGACACCAAGCTGTCCCCC 840
QY 841 tctgtactgttcaactcagacgaagattgattcaatacgaactacttgcaatcttgggt 900
Db 841 TTCTGTGACCTGTTCACCCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTGAAA 900
QY 901 aagtaactcgggttagctgtggttaaccattgggtccagctcaaggtgttgggttctggt 960
Db 901 RAGTATTACGCCATGTGTGACGTAACCCGCTCGGCCGACCCAGGGGGTGGGTACGCT 960
QY 961 aacgaattgattgttagattgactcactctccagttcgaagaccacacttactaataccac 1020
Db 961 AACGAGCTCATCGCCCGCTGACCCACTCGCTGTCCACGATGACACGAGTTCGAAGCAC 1020
QY 1021 actttggacttaaccagctacttctccattgaacttcttgaacttcttgcgtgacttctct 1080
Db 1021 ACTTTGGACTCGAGCCCGGCTACCTTTCCGCTCACTCTACTCTCTACGCGGACTTTTCG 1080
QY 1081 cagcaaacactatgtttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1140
Db 1081 CATGACAACGGCATCATCTCCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1140
QY 1141 ttgtactacttctgttgaactctattgaagaaactgacggttactctgttcttggact 1200

Db 1141 CTATCTACCACGACCGTGGAGAAATATACCCAGACAGATGATCTCGTCTCTGTTGGAG 1200
QY 1201 gtccattcgtcgttagagcttgaattgaatgacgaatgacgaatgacgaatgacgaatgacga 1260
Db 1201 GTTCCGTTTGTCTCGGCTTTGACGTCGAGATGATGCACTGTACGCGGAGCAGGAGCGG 1260
QY 1261 ttggttagagtttggtaaacacagagattgttccattgcaacggttgggtgttgaacag 1320
Db 1261 CTGGTCCGTTCTTGTATGATCGGTTGCCGCTGCATGGGTGTCGCGTTGATGCT 1320
QY 1321 ttggttagatgtaagagacagactcgttgaaggttcttcttctcgttagatctggtggt 1380
Db 1321 TTGGGAGATGTACCCGGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGT 1380
QY 1381 aactgggaagaatgttctcctta 1403
Db 1381 GATTGGCGAGTGTTCCTTA 1403

RESULT 4
US-08-693-709-1
; Sequence 1, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OIJEN, ALBERT J.J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1401
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: 70...1401
; OTHER INFORMATION:

NAME/KEY: Signal Sequence
LOCATION: 1...72
OTHER INFORMATION:
US-08-693-709-1

Query Match 43.7%; Score 614.2; DB 1; Length 1404;
Best Local Similarity 64.9%; Pred. No. 4.6e-173;
Matches 910; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

```
Qy 1 atggcggttcgtcgtagctactgtccattgccacattgttgcgttccacatcccggtacc 60
Db 1 atggcggttcgtcgtagctactgtccattgccacattgttgcgttccacatcccggtacc 60
Qy 61 gcttgggttcgtcgtagctactgtccattgccacattgttgcgttccacatcccggtacc 120
Db 61 gcttgggttcgtcgtagctactgtccattgccacattgttgcgttccacatcccggtacc 120
Qy 121 tcccagaaattcttccactgttgggttacatactctccattctctcttcttgggtgacgaa 180
Db 121 tcccagaaattcttccactgttgggttacatactctccattctctcttcttgggtgacgaa 180
Qy 181 tctgtattctccagagcttccaaaggttgcgttgcgttgcgttgcgttgcgttgcgttgc 240
Db 181 tctgtattctccagagcttccaaaggttgcgttgcgttgcgttgcgttgcgttgcgttgc 240
Qy 241 agacacggtgctagatatacccaactcttctgcgtcgttaagcggtactctgcttgcgttgc 300
Db 241 agacacggtgctagatatacccaactcttctgcgtcgttaagcggtactctgcttgcgttgc 300
Qy 301 gctattcaaaagagctactgttccaaaggttaagtcacgttcttcttgaagacttacac 360
Db 301 gctattcaaaagagctactgttccaaaggttaagtcacgttcttcttgaagacttacac 360
Qy 361 taactttgggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 420
Db 361 taactttgggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 420
Qy 421 attagttctacagaagatacaaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 480
Db 421 attagttctacagaagatacaaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 480
Qy 481 tctggttctcagagcttattgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 540
Db 481 tctggttctcagagcttattgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 540
Qy 541 aagttggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 600
Db 541 aagttggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 600
Qy 601 ccagaaggtgctggttccacacacacttggaccaggttgcgttgcgttgcgttgcgttgcgttgcgttgc 660
Db 601 ccagaaggtgctggttccacacacacttggaccaggttgcgttgcgttgcgttgcgttgcgttgcgttgc 660
Qy 661 tctgaattgggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 720
Db 661 tctgaattgggtgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 720
Qy 721 gctagattggaagctcacttgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 780
Db 721 gctagattggaagctcacttgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 780
Qy 781 atggacatggttcattgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 840
Db 781 atggacatggttcattgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 840
Qy 841 tctgtgacttgcgttgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 900
Db 841 tctgtgacttgcgttgcagcagcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 900
Qy 901 aagttactaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 960
Db 901 aagttactaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 960
```

RESULT 5

US-08-419-448-33
Sequence 33, Application US/08419448
Patent No. 5863533
GENERAL INFORMATION:
APPLICANT: Robert F.M. Van Gorcom
APPLICANT: Willem Van Hartingsveldt
APPLICANT: Petrus A. Van Paridon
APPLICANT: Annemarie B. Veenstra
APPLICANT: Rudolf G.M. Luttin
APPLICANT: Gerardus Seiten
TITLE OF INVENTION: Cloning and Expression of Microbial
TITLE OF INVENTION: Phytase
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,448
FILING DATE: 10-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20026.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:

LENGTH: 1404 base pairs
TYPE: nucleic acid.
STRANDEDNESS: double
TOPOLOGY: linear.
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus ficuum (Aspergillus niger)
STRAIN: NRRL 3135
US-08-419-448-33

Query Match 43.7%; Score 614.2; DB 2; Length 1404;
Best Local Similarity 64.9%; Pred. No. 4.6e-173;
Matches 910; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

Qy 1 atggcggtgttcgtgctgactgacactcctctgacactgtgacggtgtgacaaatgt 120
Db 1 atggcggtgttcgtgctgactgacactcctctgacactgtgacggtgtgacaaatgt 120

Qy 61 gcttgggtgttcgtgacactcctctgacactgtgacggtgtgacaaatgt 120
Db 61 gcttgggtgttcgtgacactcctctgacactgtgacggtgtgacaaatgt 120

Qy 121 tcccgagaatttccactgtggtgtgacactcctctgacactgtgacggtgtgacaaatgt 180
Db 121 tcccgagaatttccactgtggtgtgacactcctctgacactgtgacggtgtgacaaatgt 180

Qy 121 ttcctcgagacttccgacttctggtgtgacactcctctgacactgtgacggtgtgacaaatgt 180
Db 121 ttcctcgagacttccgacttctggtgtgacactcctctgacactgtgacggtgtgacaaatgt 180

Qy 181 tctgctatttccagagtcgacactcctctgacactgtgacggtgtgacaaatgt 240
Db 181 tctgctatttccagagtcgacactcctctgacactgtgacggtgtgacaaatgt 240

Qy 241 agacaggtgtgacactcctctgacactgtgacggtgtgacaaatgt 300
Db 241 agacaggtgtgacactcctctgacactgtgacggtgtgacaaatgt 300

Qy 301 gctattcaaaagacgctactgcttccaaaggtgacactcctctgacactgtgacggtgtgacaaatgt 360
Db 301 gctattcaaaagacgctactgcttccaaaggtgacactcctctgacactgtgacggtgtgacaaatgt 360

Qy 361 tacacttgggtgtgacactcctctgacactgtgacggtgtgacaaatgt 420
Db 361 tacacttgggtgtgacactcctctgacactgtgacggtgtgacaaatgt 420

Qy 421 attagttctacagagatacagagcttggctgacagagcttggctgacagagcttggctgacagagct 480
Db 421 attagttctacagagatacagagcttggctgacagagcttggctgacagagcttggctgacagagct 480

Qy 481 tctggttctgacagagttattgcttctgctgacagagcttggctgacagagcttggctgacagagct 540
Db 481 tctggttctgacagagttattgcttctgctgacagagcttggctgacagagcttggctgacagagct 540

Qy 541 aagttgggtgtgacactcctctgacactgtgacggtgtgacaaatgt 600
Db 541 aagttgggtgtgacactcctctgacactgtgacggtgtgacaaatgt 600

Qy 601 ccaaaagtgctggtttacacacacitgtgacacagcttggctgacagagcttggctgacagagct 660
Db 601 ccaaaagtgctggtttacacacacitgtgacacagcttggctgacagagcttggctgacagagct 660

Qy 661 tctgaattgggtgtgacagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 720
Db 661 tctgaattgggtgtgacagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagcttgaagct 720

Qy 721 gctagattggagctgacactcctctgacactgtgacggtgtgacaaatgt 780
Db 721 gctagattggagctgacactcctctgacactgtgacggtgtgacaaatgt 780

Qy 781 atgacatgtgtcctacacacacitgtgacacagcttggctgacagagcttggctgacagagcttggct 840
Db 781 atgacatgtgtcctacacacacitgtgacacagcttggctgacagagcttggctgacagagcttggct 840

RESULT 6
US-07-923-724-7
; Sequence 7, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambossek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerström, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992

Qy 841 ttctgtgactgttctcaactcagcagaatggtattcaactacgactactgtgacttcttgggt 900
Db 841 ttctgtgactgttctcaactcagcagaatggtattcaactacgactactgtgacttcttgggt 900

Qy 901 aagttactcaggttacggtgtggttaacacattgggttcacagctcgaaggtgttgggttctt 960
Db 901 aagttactcaggttacggtgtggttaacacattgggttcacagctcgaaggtgttgggttctt 960

Qy 961 aagttactcaggttacggtgtggttaacacattgggttcacagctcgaaggtgttgggttctt 1020
Db 961 aagttactcaggttacggtgtggttaacacattgggttcacagctcgaaggtgttgggttctt 1020

Qy 1021 acttgggacttaacccagctacttcccacttgcattgaacgctacttggtagcgtgacttctct 1080
Db 1021 acttgggacttaacccagctacttcccacttgcattgaacgctacttggtagcgtgacttctct 1080

Qy 1081 cagcacaacactatggtttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1140
Db 1081 cagcacaacactatggtttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1140

Qy 1141 ttgtctactacttctgttgaaatcttattgaagaaactgacggttactctgtcttcttggact 1200
Db 1141 ttgtctactacttctgttgaaatcttattgaagaaactgacggttactctgtcttcttggact 1200

Qy 1201 gttccactcgtcgttagagcttactgtgaaatgatgcaatgtgaagctgaaagaaagaaacca 1260
Db 1201 gttccactcgtcgttagagcttactgtgaaatgatgcaatgtgaagctgaaagaaagaaacca 1260

Qy 1261 ttgggttagagtttgggttaacagagagtttccactgacaggtgttgggtgttgggtgttgggt 1320
Db 1261 ttgggttagagtttgggttaacagagagtttccactgacaggtgttgggtgttgggtgttgggt 1320

Qy 1321 ttgggttagagtttgggttaacagagagtttccactgacaggtgttgggtgttgggtgttgggt 1380
Db 1321 ttgggttagagtttgggttaacagagagtttccactgacaggtgttgggtgttgggtgttgggt 1380

Qy 1381 aactgggaagaatgttctgctta 1403
Db 1381 aactgggaagaatgttctgctta 1403

TITLE OF INVENTION: In Trichoderma
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/609,426A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/923,724
 FILING DATE: 31-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/496,155
 FILING DATE: 19-MAR-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/044,077
 FILING DATE: 29-APR-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 8610600
 FILING DATE: 30-APR-1986
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Grant E.
 REGISTRATION NUMBER: P-41,264
 REFERENCE/DOCKET NUMBER: 1050.0080001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2363 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: both
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(404...447, 550...1906)
 US-08-609-426A-7

Query Match 43.5%; Score 610.2; DB 2; Length 2363;
 Best Local Similarity 64.9%; Pred. No. 9e-172;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

QY	313	aacgtactgtcttcaagggtgaagtacgtcttcttgaagacttaacaactacacattgggt	372
DB	818	AACGTGACCACCTTTGATGAAAAATATGCTTCCTGAAGACATACAATACAGTTGGGT	877
QY	373	gctgaagacttgactccattcggtgaacacaaatgggttaactctgggtatttaagtctac	432
DB	878	GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGCTAGTCACTCCGGGATCAAGTTCATC	937
QY	433	agaagatacaaggcttggcttagaagattgttccattcatttagagcttctggttctgac	492
DB	938	CAGCATACGAATCGCTCACAGGACACATCATTCGTTTCATCCGATCTCTGGCTCCAGC	997
QY	493	agagttattgcttctgtgaaagtctcaatgaaggtttccaatctgtctaagttggctgac	552
DB	998	CGGTGATCGCTCCGGCGAGAAATTCATTGAGGGCTTCCAGAGCACCAGCTGAAGGAT	1057
QY	553	caaggtgtaacccacacacaaagctctccagttataaagttattattccagaagtgct	612
DB	1058	CCTCGTGCACCGCGGGCCAATCGTCGCCCAAGATCGAGGTGGTTCATTTCGAGGCCAGC	1117
QY	613	ggttaacaacaacacttggaccacggttggtagctgtcttcgaagaatctgaattgggt	672
DB	1118	TCATCCACAACACACTCTGACGCCAGGCACCTGCACCTGTCTTTGAGAGCAGCAATTGGCC	1177
QY	673	gacgagcttgaagcttaactcactgtgttttgcgtccaccaattagagctagattgaa	732
DB	1178	GATACCGTCGAAGCCAAATTTACCGCCACGTTTCGCCCTCCATTCGTCACAGCTGGAG	1237
QY	733	gctcaattgccagggtttaactgactgaagagagctgttgaacttgatgagaaatggt	792
DB	1238	AACGACCTGTCTGGCGTGACTCTCACAGACACAGAGTGACCTACCTCATGGACATGTC	1297
QY	793	ccattcgacactgtgttagaactcttcgactcactcaattgtctccattctctgtgacttg	852
DB	1298	TCCTTCGACACCATCTCCACAGCACCGTTCGACACCAACCTGTCCCTCTCTGTGACCTG	1357
QY	853	ttaactcaaggaatggattcaactacgactacttgcaatcttttgggttaagtaactaggt	912
DB	1358	TTCAACCCATGACGAATGGATCCACTACGACTACCTCCAGTCCCTGAAAAAATACTACGC	1417
QY	913	tacggtgctgtaacccattgggtccagctcaaggtgttgggtttgttgaacgaattgatt	972
DB	1418	CATGGCGCAGGTAAACCCGCTCGCGCGACCCAGGGCGTCGGCTACGCTAACGAGCTCATC	1477
QY	973	gctagattgactcaactctccagttcaagaccacactcttactaaaccacactttggactct	1032
DB	1478	GCCGCTCTACCCACTCGCCTGTCCACAGATGACACACAGCTCCACACACACCTTGGACTG	1537
QY	1033	aaccagctactttccattgaacgctactttgtacgtgacttctctcaacacacact	1092
DB	1538	AACCCAGCTACCTTCCCGCTCAACTCTACTCTCTAGCGGAGCTTTTCCACAGATAACGC	1597
QY	1093	atggtttctattttctgctttgggtttgtacaaaggttaactaaagcattgtctactact	1152
DB	1598	ATCATCTCTATCTCTTGTCTTGGGTCTGTACAGGGCAGCTAAGCCGCTGTCTACACAG	1657
QY	1153	tctgttgaactctattgaagaactgacggttactctgtcttcttggactgttccattcgt	1212
DB	1658	ACCGTGAGAAATATACCACGACAGATGGTTCGCTCTGCTTGGAGCGTTCGGTTGCT	1717
QY	1213	gctagagcttaacttgaatgatgcaatgtaagctgaagaagaaacccattggttagatt	1272
DB	1718	TCGCGCTGTGTACGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1777
QY	1273	tgtgttaagacagagatttccattgacggttgggttgggttgggttgggttgggttgggt	1332
DB	1778	TTGGTTAATGATCGCGTGTTCGCGCTGCATGAGGTGTCCAAATGATGATGATGATGATGAT	1837
QY	1333	aagagagagacttctgaaaggtttgtctctgctagatctggttgggttaactgggaagaa	1392
DB	1838	ACCCGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGGGCGAG	1897
QY	1393	tggttcgctta 1403	

Db 864 COTCGTGCCAGCCGCGCAATCTGCGCCAAAGATCGAGTGGTCTATTTCCGAGCCAGC 923
Qy 613 ggttaacaacacactttgacacacggtttgtgtactgttttcgaagaatctgaattggt 672
Db 924 TCATCCACACACTCTCGACCCAGCCAGCTGCTCTCGAAGACAGCGAATTGGCC 983
Qy 673 gacgagttgaagctaaacttactgtttttcgtccaccaatagagctagattgaa 732
Db 984 GATACCGTCGAAGCAATTTACCCGCCAGTTCGTCCCTCCATTCGTCAAGCTCTGGAG 1043
Qy 733 gctcaactgcaggtgttaacttgaactgacgaagacgttgaatcttatgagacatgtgt 792
Db 1044 AACGACCTGTCCGCTGACTCTACACACAGAGAGTACCTCATGAGCATGTGC 1103
Qy 793 ccattcgacactgttgtagaactctacactactcaactgtctccattctgtgacttg 852
Db 1104 TCCTTCGACACCATCTCCACAGCAGCGTCGACACCAAGCTGTCCCTTCTGTGACCTG 1163
Qy 853 ttactcagcgaatgattcaatacagactacttgcgaactcttttgggtaagtactacgt 912
Db 1164 TTCACCCATGAGGAATGATCAACTACGACTACCTCCAGTCTTGAAGATTTACGGC 1223
Qy 913 tacggtgtgttaacccacttgggtcgaactgaagtggtgttgcgttaacgaattgatt 972
Db 1224 CATGTGCGAGTTAACCGCTCGCGCCGACCCAGGGCGTGGCTAGCTAACGAGCTCATC 1283
Qy 973 gctagatgactcacttccagttcaaacacacacttctactaacacacacttggactct 1032
Db 1284 GCCGCTTGACCACTCGCTGTCACGATGACACCAAGTTCCACACCACTTTGGACATCG 1343
Qy 1033 aaccagctacttctcccaattgaacacttctgtactgttctctcagcacaacact 1092
Db 1344 AGCCCGGCTACTCTCCGCTCAACTCTACTCTCTACGGGACTTTTCGATGACACGGC 1403
Qy 1093 atggtttctatttctcgtcttgggttgggttgaacacgtaactgaacattgtctactact 1152
Db 1404 ATCATCTCCATCTCTGCTGCTTAGTCTGTACACGGCGCTAAGCGGTATCTACCAAG 1463
Qy 1153 tctgttgaactattgaagaactacggttactctgttctgtgactgttccattcgtct 1212
Db 1464 ACCGTGGAGAAATATCACCCAGACAGATGATTTCTGCTGTGACGGTTCGGTTGCT 1523
Qy 1213 gctagagcttctgaaatgatgcaatgtgaagctgaagaaagaccattggttagagtt 1272
Db 1524 TCGCGTTTGTAGTCGAGATGATGAGTGTACAGCGGACGACGAGCGCTGTCGCTGTC 1583
Qy 1273 ttggttaacgacagattgttccattgcaacggttgtgtgttgacaagtgggttagatgt 1332
Db 1584 TTGGTTAATGATCGCGTTGTCCTCCGCTGCATGGGTGTCCGGTTGATGCTTTGGGAGATGT 1643
Qy 1333 aagagagacgacttctgtgaaggttcttctcgttagatctcgttgggttaactgggaagaa 1392
Db 1644 ACCCGGATAGCTTGTGTGAGGGGGTTGAGCTTTGTAGATCTGGGGGTGATTGGCGGAG 1703
Qy 1393 tgtttcgtcta 1403
Db 1704 TGTGTTGCTTA 1714

RESULT 10

US-08-419-448-31

; Sequence 31, Application US/08419448

; Patent No. 5863533

; GENERAL INFORMATION:

; APPLICANT: Robert F.M. Van Gorkom

; APPLICANT: Willem Van Hartingsveldt

; APPLICANT: Petrus A. Van Paridon

; APPLICANT: Annemarie E. Veenstra

; APPLICANT: Rudolf G.M. Luttin

; APPLICANT: Gerardus Seiten

; TITLE OF INVENTION: Cloning and Expression of Microbial

; TITLE OF INVENTION: Phytase

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; IMMEDIATE SOURCE:
; LIBRARY: lambda AF
; CLONE: pAF2-3, pAF2-6, pAF2-7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 210..253
; FEATURE:
; NAME/KEY: intron
; LOCATION: 254..355
; FEATURE:
; NAME/KEY: exon
; LOCATION: 356..1715
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(210..253, 356..1715)
; OTHER INFORMATION: /codon_start=210
; -OTHER INFORMATION: /product="Phytase"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 210..380
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 381..1712
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function="inositol phosphate
; OTHER INFORMATION: /phosphate
; OTHER INFORMATION: /product="Phytase"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; US-08-419-448-31

Query Match

43.5%; Score 610.2; DB 2; Length 6756;

Best Local Similarity 64.9%; Pred. No. 1.5e-171;

Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgtactgtccattgcccacttctcgtgttcacatccggtaccgcttgggtcct 72

Db 324 GTGGGACTACTGATCGCTGACTATCTGTGAGTACCTCCGAGTCCGAGTCCCGGCC 383

```
Qy 73 cgtggtaactctcaactcttctgtgacactgtgtgacggtgttaccatgtttccacagaatt 132
Db 384 TCGAGAAATCAATCAGTTCGATACGGTCGATCAGGGGTATCAATGCTTCCGAGACT 443
Qy 133 tctcaactgtgggttacatactctccactctctctctctctctctctctctctctctctct 192
Db 444 TCGCATCTTTGGGGTCAATACGACACCGTCTCTCTCTCGCAACGAAATCGCTCATCTCC 503
Qy 193 ccagacgttccaaaaggtgttagatctactcttctctctctctctctctctctctctctct 252
Db 504 CCGAGGTGCCCGCGGATGACAGTCACTTTCTGCTCAGGTCTCTCTCTCTCTCTCTCTCTCT 563
Qy 253 agatacccaactctctctgctcttaaggcgtaactctgtctgtttgattgaagctattcaaa 312
Db 564 CGGTATCCGACCGACTCCAAAGGCGAAGAAATACCTCGCTCTCTCTCTCTCTCTCTCTCTCT 623
Qy 313 aacgtactgtcttcaagggttaagtacgtctctctctctctctctctctctctctctctct 372
Db 624 AACGGGACACACTTTGACGGGAAATATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 683
Qy 373 gctgacactgactccactctggttaacacaaatggttaactctggttatttaagtctac 432
Db 684 GCAGATGACCTGACTTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGCATCAAGTCTCTAC 743
Qy 433 agaagatacaagcgttctgctagaagattgttccattcattagagctctctggttctgac 492
Db 744 CAGCGGTAGGAATCGCTCACAGGAACATCGTTCATTCATCCGATCTCTGGCTCCAGC 803
Qy 493 agagtattgtctctgaaagtctcattgaaggttcccaactctgtaagtgtgctgac 552
Db 804 CCGGTGATCGCTCCGGAAGAAATTCATCGAGGCTCTCCAGAGCACCAAGTGAAGAT 863
Qy 553 ccaggtgttaacccacacaaagctctccagttatttaagttatttccagaagtgct 612
Db 864 CCTGTGCCCCAGCCCGCAATCGTGCACAGATCGACGTGCTCATTCGAGGCGCCAGC 923
Qy 613 ggttacaacacactttggacacggttctgtactgtctctctctctctctctctctctctct 672
Db 924 TCATCAACACACACTCTCGACCCAGGACCTGCACCTGCTTCGAAGACAGCGAATTGGCC 983
Qy 673 gacgacgtgaagctaaactctactgctgttctctctctctctctctctctctctctctct 732
Db 984 GATACCGTCGAAGCGAATTTACCCGCCACGTTCTGCTCCCTCCATTCGTCAAGCTCTGGAG 1043
Qy 733 gctcaactgcccaggtgttaacttgaactgaacgaacggtgttgaacttgaacatgtgt 792
Db 1044 AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTGACCTACCTCATGGACATGTGC 1103
Qy 793 ccattgacactgttctgtagaactctgacgctactcaaatgtctccattctgtgacttg 852
Db 1104 TCCTTCGACACCATCTCCACGACCGCTCGACACCAAGCTGTCCCTCTCTGTGACCTG 1163
Qy 853 ttcactcaacgaagaatgatcaatacactacttgaacttgaacttgaacttgaacttgaact 912
Db 1164 TTCACCATGACGAATGATCAACTTACGACTACCTTCCAGTCTCTCTCTCTCTCTCTCTCTCT 1223
Qy 913 tacggtgtctglaacccactgtggttccagctcaaggtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1224 CATGGTGCAGTAACCCCTCGCCCGGACCCAGGCGCTGCGCTAGCTTAACGAGCTCATC 1283
Qy 973 gctagatgactcactctccagtttcaagaccacactcttacttaacacacacttggactct 1032
Db 1284 GCCGTCTGACCCACTCGCTCTGTCACGATGACACCAAGTTCACACCACTTTGGACTCG 1343
Qy 1033 aaccagactatttcccaattgaagcactacttgttaacgtgacttctctctcaacacact 1092
Db 1344 AGCCCGGTACTCTTCCGCTCAACTCTACTCTACGCGGACTTTTTCGATGACAACGCG 1403
Qy 1093 atggttctatttctctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1152
Db 1404 ATCATCTCAATCTCTTTCGCTTTAGGCTGTGACACGCGCACTAAGCCGCTATCTACCACG 1463
```

```
Qy 1153 tctgttgaactctattgaagaaactgacggttactctgctctctctgactgttccattcgt 1212
Db 1464 ACCGTGGAGAAATATCACCCAGACAGATGGATCTCTGCTGCTTGGACGGTTCCTGTTGCT 1523
Qy 1213 gctagagcttactgttgaatgatcaatgtgaagctgaaagaaagaaacccattggttagagt 1272
Db 1524 TCGCGCTTTGTACGTGAGATGATGATGCTGACGGCGGAGCAGGCGCTGCTCGGTGTC 1583
Qy 1273 ttggttaacgacagaggttcttccattgacagcttctgtgttgacaaagtgggttagagt 1332
Db 1584 TTGGTTAATGATCGCGTTGCTCCGCTGATGGGTGCTCCGTTGATGCTTTGGGGAGATGT 1643
Qy 1333 aagagagacgactctgttgaaggttcttctctgctgaatctgtgtgtaactgggaagaa 1392
Db 1644 ACCCGGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGTCAATTGGGGGAG 1703
Qy 1393 tgttctgctta 1403
Db 1704 TGTCTTGTCTTA 1714
```

RESULT 11

```
US-09-155-855-5
; Sequence 5, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1)..(45)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(45)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (157)..(183)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (184)..(1512)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)..(1512)
US-09-155-855-5
```

Query Match 42.18; Score 591.6; DB 3; Length 1515;
Best Local Similarity 64.6%; Pred. No. 2.5e-166;
Matches 882; Conservative 0; Mismatches 484; Indels 0; Gaps 0;

```
Qy 38 tgtctcgttccacatccggttaccgcttggttctctctgttgtaactctcaactcttgata 97
Db 149 tgcacagagtccactccgactggcagccccgcctcctcctgagaaatccacttgagata 208
Qy 98 ctgttgacggtgttaccatgtttccacagaattctcacttgggtgtacatactctc 157
```

Db 209 cggctgacaggggtatcaatgtctctcggagacttcgcatcttggggcccaatacgcgc 268
Qy 158 cattctctcttggctgacgaatctgtatttctccagaagttccaaagggttgtagag 217
Db 269 cgttctctctcggcaacaatacggcatctccctctgattctctcgggatgccatg 328
Qy 218 ttacttctgctcaagtgttcttagacacggtgtctagatacccaactctctctgctcta 277
Db 329 tcaatttgcggcaggctctctccgcacatgagacggtatccgacgactccaaaggcca 388
Qy 278 aggcgtactctgttctgattgaagctatccaaagaacgctactcttctcaagggttaagt 337
Db 389 agaaatactcgcctctcatcagggaggtccagcaacgacgacacaccttgcgggggaaat 448
Qy 338 acgcttcttgaagaacttacaactacacttgggtgtctgagacgacttgactccattcggtg 397
Db 449 atgctctctgaagacatacaactacagcctggcgcgagatgactgactccctcggag 508
Qy 398 aacaacaaatggttaactctgttataagttctacagaagatacaaggcttcttggctagaa 457
Db 509 agcaggagctggtcaactccggcgctcaagttctacacgacgatacgaatcgctcacaagaa 568
Qy 458 agattgttccattcattagagctctgttctgtacagagttattgtctctgtctgaaaagt 517
Db 569 acattgtccgctctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 628
Qy 518 tcaattgaaggttccaaatctgtaagtgtgtgacccaggtgtcactaccacacacagctt 577
Db 629 tcatcgaggtctccagacactaagctgaaggtatctctgtccagccggcccaatgtg 688
Qy 578 ctccaggtatttaacgttattatccagaaggtgctgtgttacaacacacacttggaccacg 637
Db 689 cggccaaagatcgagctgtcatttccagaggccagcacatcccaacacactctcgatccgg 748
Qy 638 gtttgtactgtcttccgaagaatctgaattgggtgacgacgttgaagtaacttcaactg 697
Db 749 gcacctgcacctgttttccgaagtagcgaattggcgatgacatcgaaagcgaatttcaacg 808
Qy 698 ctgtttcgtcccaacattagatagattggaagctcaattgcccaggtgttaacttga 757
Db 809 ccactgtctcctccatctcagctcagctcagctcagctcagctcagctcagctcagctc 868
Qy 758 ctgacgaagacgtgttaactgtatgagacgtgtccattcgacactgtgtgtagaactt 817
Db 869 cggacacagaagtgacactaccctcagatggaatgtctctcagacacactcaccacagca 928
Qy 818 ctgacgctactcaattgtctcattctgtgactgtgttcaactcacgacgaatggattcaat 877
Db 929 ccgtcgacacacagctgtccctctctgtgacctgttccacctgaagaatggatcaact 988
Qy 878 acgactacttgaatttgggtgaagtagactacggttaccggtggtggttaaccacattgggtc 937
Db 989 acgactactcagctccttccgaacataactacggtcagctcagctcagctcagctcagctc 1048
Qy 938 cagctcaaggtgtgtgttctgtaacgaattgattgctagattgactcactctccagttc 997
Db 1049 cagccaggcgtcgtcagctcagctcagctcagctcagctcagctcagctcagctcagctc 1108
Qy 998 agacacacacttctactaaccacacttggactcttaaccacgactacttcccaatgaacg 1057
Db 1109 acgatgacacacagctcccaacacacttggactcccaaccggctacttctcccgctcaact 1168
Qy 1058 ctacttgtacgctgactctctcagacacacactatggtttctatttctctcgtttgg 1117
Db 1169 ccactctctatcgactctcttgcgatgaacggcatctctctatctctctcttctgtttgg 1228
Qy 1118 gtttgtacacggttactaagcattgtctactacttctgttgaattattgaagaactg 1177
Db 1229 gctgtacacacgacacacgctgtcttccacgacccgagagaatacaccacagacg 1288
Qy 1178 acggttactctctcttctgactgttccatctcgtgctagacttactgtaaatgatgc 1237
Db 1289 atgggttctcactcgtcggcgggttctctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1348

Qy 1238 aatgtgaagctgaaaggaagacacattggttagagtttgggttaacagacagagttgttccat 1297
Db 1349 aatgcccagctccagcagcagcagccttggctccggtgtcttgggttaataatgactggtgttccgc 1408
Qy 1298 tgcacggttgggtgttgcacaaagttgggttagatgtatgaagacgactcgtttgaaggtt 1357
Db 1409 tgcagtgctgcgggtgtgatgttgggaagatgtacgcgggagtagcttctgaaaggggt 1468
Qy 1358 tgtcttgcgtagatcgtgtgttaactgggaagaatgtttcgttta 1403
Db 1469 tgaogttccagatcgtggtgagtggtggggggaggtgttctcgtta 1514
RESULT 12
US-09-155-855-4
; Sequence 4, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1332)
US-09-155-855-4

Query Match 42.1%; Score 590.6; DB 3; Length 1332;

Best Local Similarity 65.5%; Pred. No. 4, 6e-166;
Matches 863; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

Qy 87 ctctgtgacactgttgacggtgttaccatgtttccagaaaattctcaacttggggg 146
Db 15 caattgacgatacggctcgatcaggggtatcaatgtcttcggagacttcgacttggggg 74
Qy 147 tacatactctccattctctcttctgtgctgacgaatctgtatttctccagacgttccaaa 206
Db 75 ccaatacgcgcgttcttctctggtgcaacaaatcgccatctccctcgtatgtctgc 134
Qy 207 ggggtgtagaattacttctgttcaagttttgttagacacggtgttagatacccaacttc 266
Db 135 cggatgccatgtcacttctgcacaggttctctcccgatggagcgggtatccgacga 194
Qy 267 ttctgcgttaaggcgtactctgtcttggatgaagctattcaaaaagacgactactgttt 326
Db 195 ctccaaggcagaagaataactcgcctctcgcagagatccacagacagcgcgaacactt 254
Qy 327 caagggtgaagtaagcttcttgggaagacttacaactacacttgggtgtcgtacgacttgac 386
Db 255 cgagggggaataatgccttctcgtgaacatacaactacagctcgtggcggtgactgtac 314
Qy 387 tccattcgtgaaacacaaatggttaactctgttatttaagtttctacagaagatacaagc 446
Db 315 tcccttcgagagcagaggtgttcactcccgctcaagttctaccacgacgaacatc 374
Qy 447 ttgggctagaagaatgtttccattcattagagcttctgtgtctgacagagatttggcttc 506

```

Db 375 gctcacaagaacatgtccgttccatcgtactcagctccagccggtgattgctc 434
QY 507 tgcgtgaagttcattgaaggtttccatctcgtcgaagttggtgacccaggtgctaaacc 566
Db 435 tggcaataaattcagcaggtccagagcagcagcagcagcagcagcagcagcagcagc 494
QY 567 acaccaagcttccagttatcagcagcagcagcagcagcagcagcagcagcagcagc 526
Db 495 cggccaatcgtcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 554
QY 627 ttggaccacaggtttgtactgctttcgaagaatcgaatgggtgacgacggttgagc 686
Db 555 tctgcacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 614
QY 687 taactcactgctgtttcgtccacacacacacacacacacacacacacacacacacac 746
Db 615 caatttcacgcacagcttgcctccctccatcgtcagcagcagcagcagcagcagcagc 674
QY 747 tgttaacttgactgacgacgagcagcagcagcagcagcagcagcagcagcagcagcagc 806
Db 675 cgtgtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 734
QY 807 tgtagaacttctgacgctactcaatgtctccattctgtgactgttcaactcacgacga 866
Db 735 ctccaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 794
QY 867 atgattcaatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 926
Db 795 atggatcaactacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 854
QY 927 cccattgggtccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 986
Db 855 ccgcgtcgccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 914
QY 987 ctctccagttcaagaccacacttctactaacacacacacttctgactctaacacacagcagcagc 1046
Db 915 ctgcgtctccacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 974
QY 1047 cccattgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1106
Db 975 ccgcgtcaactccactctctctctctctctctctctctctctctctctctctctctctct 1034
QY 1107 ctctcgtttgggtttgtaacacgctactaacgacattgtctactactctctgtgactctat 1166
Db 1035 ctttgcgtttgggtttgtaacacgctactaacgacattgtctactactctctgtgactctat 1094
QY 1167 tgaagaactgacggttactctgttcttggactgttccattcgtcgtcgtcgtcgtcgtcgt 1226
Db 1095 caccacagaccgaggttctcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1154
QY 1227 tgaatgatgaatgtgaagctgaaaggaagacacattggttagaggtttggttaacgacag 1286
Db 1155 cgagatgatgaatgccagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1214
QY 1287 agttgtccatgacggttgcgtggtgtgacaaagttggttagatgagagagacgactt 1346
Db 1215 tgttgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1274
QY 1347 cgttgaaggtttgttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1403
Db 1275 cgtgaaggggttgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1331

```

RESULT 13

US-08-819-825-1

Sequence 1, Application US/08819825

Patent No. 5866118

GENERAL INFORMATION:

APPLICANT: Berka, Randy M.

APPLICANT: Ray, Michael W.

APPLICANT: Klotz, Alan V.

TITLE OF INVENTION: Polypeptides Having Phytase Activity

```

; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5

```

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/819,825

FILING DATE: 18-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4758.200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2200 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-819-825-1

Query Match 24.6%; Score 346; DB 2: Length 2200;

Best Local Similarity 57.2%; Pred. No. 3 le-93;

Matches 753; Conservative 0; Mismatches 515; Indels 48; Gaps 5;

```

QY 127 gaaattctcaactgtgggtacatactctccattcttcttggctgacgaattct 186
Db 607 GACATTCCCGCCCACTGGGCCAGTACTCCCTTCTCTCGGTGGCGAGGCTCTGAA 666
QY 187 attctccagacgttccaaaggttagagttacttctgttcaagtttctgtctagaac 246
Db 667 ATCTCGCTCGGTGGCCCAAGGGCTGTCTGTCGAGTTTGTGCAGGTGTCTCCGGCAC 726
QY 247 ggtgctagataccacacttctctgcttaagcgtactctgcttctgattgaagctatt 306
Db 727 GGAGTCGGTATCTCTACTGCTCACAAGAGTGAAGTCTACGCCGAGTTGCTTCAAGGATC 786
QY 307 caaagaacgctactgctttcaagggtaagtagcgtttcttgaagacttacactacact 366
Db 787 CAGGACACTCGCAGCGAGTTCAAGGGCGATTTCCTTTCCTTCGAGACTATGCTATCAT 846
QY 367 ttgggtgctgacgactgactccattcgttggaacacaaatggttaactctggttaag 426
Db 847 CTGGTGGCCGATTAATTGACGCGCTTTGGCGAGGACAGATGATGGAATCGGGCGCCAG 906
QY 427 ttctacagaagatcacaggctttgctgtagaagattgttccattcatttagagctctgct 486
Db 907 TTTTACCACCGGTATCGTGAGCAGGCGCCGAGAGATTGTGCCAATTTGTGCTGGCGAGGC 966
QY 487 tctgacagagttattgcttctctgctgaaagtctcattgaaggtttccaatctgtaagttg 546
Db 967 TCCGCGCGAGTCTATTCGTCGGCAGAGATTCTTCAACCCGGGATTCAGGATCCCAAGAC 1026
QY 547 gctgacccaggtgctaaacccacacacagctctccagttatttaacgattatttccagaa 606
Db 1027 CGGGATCCCGAGGTCGAACAAAGGACAGGAGCCCTGTGATCAACGCTGATCATTTCCGAA 1086
QY 607 ggtgctggttaacacacacactttggaccacggtttgtgtactgctttcgaagaactgaa 666
Db 1087 GAAACTGGCAGTAACAATACTCTGGATGGGCTGACGTGCCCGCGCGCGAGGA---GGCA 1143

```



```

Db 1204 ATCACTAAACACATGCGGGGTGTGAACCTCACTTGGAGGATGTCOCGTGTTCATGGAT 1263
QY 787 atgtgtccattcacactgt-----tgtagaactlctgacgtactcaattgtctcca 840
Db 1264 CTTTGCCGTTTACACAGGTGGGCTCCGACCCAGTCTTTTCCAGGGAGCTCTCTCGG 1323
QY 841 tctgtgactgttctactcaacgacgaatgagttcaatacgaactacttgcactctttgggt 900
Db 1324 TTTTGTCACTTGTTCACGGCCGACGATGGATGGCTTACGATTAATACTACACCTCGAC 1383
QY 901 aagtaactcaggttacgtgtcgtgtaaccatgggtccagctcaaggtgtggttggtt 960
Db 1384 AAATACTACACGACCCGCGGCGACGCAATTTGGCCCGTCCCGCGGCTCGGGTCTGTC 1443
QY 961 aacgaattgattgttagattgac---tcaactctccagttcaagaccacacacttctactaac 1017
Db 1444 AAGGAGCTGATTGCGGTATGACGGGAATCTCCGTCAGGACCCACACACATCAAC 1503
QY 1018 cacactttggaacttaaccagactacttcccatgaaacgtacttctgacgtgacttc 1077
Db 1504 CACACTCTCGATGACAACCCGGAACCTTCCCGTTGGAGCTGTCTCTACGACACTTT 1563
QY 1078 tctcagacacacactatgtttctattttcttctgctttggtttgtacaaagggtactaa 1137
Db 1564 TCGCAGCACAAACACCATGACGGGCTATCTTCCGCAATGGGCGCTGTACAACGGCACAAAG 1623
QY 1138 ccattgtactacttctgttgaatct-----attgaagaaactgacgttactctgct 1191
Db 1624 CCGCTGTGACGCTCAAGATTACGCTCCGAGGGGTGACGAGCGGATGGATATGCGGCA 1683
QY 1192 tcttggaactgttccatctgctgtcgtagacttacgttgaaatgtaaatgtgaagctgaa 1251
Db 1684 TCGTGGACGGTCCCGTTCGACGAGGGCGTATGTGGAGTTGCTGCGATGTGAGACGGAA 1743
QY 1252 aag-----gaaccattggttagagtttgggttaac 1281
Db 1744 AGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1803
QY 1282 gacagagtgttccatgacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
Db 1804 GATCGGGTGTGCGGCTGCATGTGTCGGGTGTGATCGATGGGAGGAGGAGGAGGAGGAGG 1863
QY 1342 gactcgttgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1397
Db 1864 GAGTGGATTAAGGAGTCACTTGTTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1919

```

RESULT 15

US-09-221-654-1

; Sequence 1, Application US/09221654

; Patent No. 6054306

; GENERAL INFORMATION:

; APPLICANT: Lassen, Soren

; APPLICANT: Bech, Lisbeth

; APPLICANT: Fuglsang, Claus

; APPLICANT: Ohmann, Anders

; APPLICANT: Breinholt, Jens

; APPLICANT: Ostergaard, Peter

; TITLE OF INVENTION: Peniophora Phytase

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6054306 No. 6054306disk of No. 6054306th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/221,654
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/989,358
; FILING DATE: 12-DEC-1997
; APPLICATION NUMBER: 0529/97
; FILING DATE: 07-MAY-1997
; APPLICATION NUMBER: 60/046,081
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 5101.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-221-654-1

```

```

Query Match 9.4%; Score 131.4; DB 3; Length 1320;
Best Local Similarity 56.9%; Pred. No. 1.5e-29;
Matches 268; Conservative 0; Mismatches 191; Indels 12; Gaps 1;

```

```

QY 663 tgaattgggtgacgacggtgaaagtaacttaactgctgttttctccaccataagagc 722
Db 594 TGAAGTGGATGGTGCAGCAATCCACAACGTTGGTGGGGTCTTTGGCCGCAACATCACCG 653
QY 723 tagattggaagctcacttgccaggtgttaacttgactgacgacgaagcgttgtaacttgat 782
Db 654 GCGATTGAACGCTGTGCGCCGAGTGCCAACTCTCAGACAGCGAGCGCTCCTACTCTCAT 713
QY 783 ggaatgtgtccattgcacactgtgtgtagaaacttctgacgttactcaattgtctccatt 842
Db 714 GGATATGTGCCGTTTCGACACTCTCAG-----CTCCGGGAACGCCAGCCCTT 761
QY 843 ctgtgactgttcaactcagacgaatggaattcaatacgaactacttgcaatctttgggtaa 902
Db 762 CTGTGACCTATTTCACCGGAGGAGTATGTGTCGTACGAGTACTACTATGACCTCGACAA 821
QY 903 gtactacggttacggtgtgtgtaaccattgggtccagctcaaggtgtgtgtgtgtgttaa 962
Db 822 GTACTATGGCAGGGCCCGGGAACGCTCTCGTCTGTCCAGGGCGGTGGATAGTCAAA 881
QY 963 cgaattgattgttagattgactcaactctccagttcaagaccacacttctacttaaccac 1022
Db 882 TGAGTGTGTTGACGCTTGACCGGCCAAGCCGTTTCAGACGAGACGAGACGAGCAACCGCAC 941
QY 1023 ttgtaacttaaccagactacttcccatgaaacgtacttgtgacgtgacttctctca 1082
Db 942 GCTGACAGCGACCTTGAACATCCCGTGAACCTCCCGTGAACCTTCTACGCGGACTTCTCGCA 1001
QY 1083 cgacaacactatggtttctattttcttcgctttggtgtgtgtgtgtgtgtgtgtgtgtgt 1133
Db 1002 TGATAACACCATGGTGGCCATCTTTTGGCGGCGCTCGGGCTCTTCAACGCCAC 1052

```

```

Search completed: October 26, 2001, 15:20:39
Job time: 595 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:29 ; Search time 48.78 Seconds
(without alignments) 729.264 Million cell upda

```

Title: US-09-488-265-31
Perfect score: 2468
Sequence: 1 MGVFVLLSIATLFGTSGT.....DFVEGLSFARSGGNWEECF 467

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	%			DB	ID	Description
			Match	Length	DB			
1	1852	75.0	467	1	JN0889	3-phytase (EC 3.1.1.		
2	1848	74.9	467	1	JN0656	3-phytase (EC 3.1.1.		
3	1787	72.4	441	1	JN0482	3-phytase (EC 3.1.1.		
4	405	16.4	467	1	PBYC	acid phosphatase (
5	394	16.0	467	2	S53476	acid phosphatase (
6	394	16.0	467	2	S48996	acid phosphatase (
7	391	15.8	467	1	PBYCC	acid phosphatase (
8	369	15.0	468	2	S52495	acid phosphatase h		
9	360.5	14.6	479	1	JN0715	3-phytase (EC 3.1.1.		
10	354.5	14.4	479	1	JN0890	acid phosphatase (
11	350	14.2	468	2	JC4285	acid phosphatase (
12	328	13.3	453	1	A25326	acid phosphatase (
13	311.5	12.6	463	2	S19929	acid phosphatase (
14	303.5	12.3	463	2	S14119	thiamin-repressibl		
15	153	6.2	465	2	JE0369	acid phosphatase (
16	149.5	6.1	468	2	A86233	histidine acid pho		
17	142	5.8	755	2	T19118	hypothetical prote		
18	130	5.3	449	2	T15933	acid phosphatase h		
19	127.5	5.2	423	2	A33395	hypothetical prote		
20	118.5	4.8	421	2	S14742	acid phosphatase (
21	118	4.8	380	2	T16883	acid phosphatase (
22	115.5	4.7	381	2	JH0152	hypothetical prote		
23	114.5	4.6	423	1	S06167	acid phosphatase (
24	109.5	4.4	730	2	T16455	acid phosphatase (
25	107	4.3	954	2	T21241	hypothetical prote		
26	106.5	4.3	602	2	A36715	hypothetical prote		
27	103	4.2	452	2	T20556	exo-poly-alpha-gal		
28	102.5	4.2	1225	2	T16346	hypothetical prote		
29	101.5	4.1	413	2	JW0087	glucose-1-phosphat		

ALIGNMENTS

RESULT 1

JN0889

3-phytase (EC 3.1.3.8) A precursor - Aspergillus awamori

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein

C:Species: Aspergillus awamori

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0889

R:Piddington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0889

A:Molecule type: DNA

A:Residues: 1-467 <PTD>

A:Cross-references: GB:L02421; NID:gl66518; PIDN:AAA16898.1; PTD:gl66519

A:Experimental source: strain ALK0243

A:Note: part of the sequence, including the amino end of the mature protein, was conf

C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic

C:Genetics:

A:Gene: phyA

A:Introns: 15/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-467/Product: 3-phytase A #status experimental <NAT>

F:27, 59, 105, 120, 207, 230, 339, 352, 376, 388/Binding site: carbohydrate (Asn) (covalent) #

F:81, 361/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match	75.0%	Score	1852	DB 1	Length	467			
Best Local Similarity	74.3%	Pred. No.	1.1e-137						
Matches	347	Conservative	42	Mismatches	78	Indels	0	Gaps	0
QY	1	MGVFVLLSIATLFSGTSGTALGPRGNHSCDVTDDGYQCQCFPELSHLWGTYSPFFSLADE	60						
Db	1	MGVSALLFLYLALGVTSGLVAPASRNQSTCDTVDQGYQCFTSTHLSWGQYAPFFSLANE	60						
QY	61	SAISDPVPGKCRVTPEVQLSRHGARYPTSSASKAYSALIEAIQKNATPKGYAFUKTVN	120						
Db	61	SAISDPVPAGCHVTFQAULSRHGARYPTESGKKYSALIEIQONTTDDGYAFUKTVN	120						
QY	121	YTLGADDLTPGGEQMWNSGKTFYRRYKALARKIVPTIRASGSDRVASAEKFIEGFOSA	180						
Db	121	YSLGADDLTPGGEQLVNSGTFKYRYESTRIILFPIRSSGSRVSIASGEKFIEGFQST	180						
QY	181	KIADPGANPHOASPVNVIPEEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR	240						
Db	181	KLKDRPAQGGSSPKIDVVIIEASSSNTLDPGCTVFDESELADTVEANFTATFAPSIR	240						
QY	241	ARLEAHLPGVNLTDDEVNLMDCPFTVARTSDATQLSPFCDLFTHDSEWQYDYLQSLG	300						

Db 241 ORLENDLSGVTLTDEVTYLLMDMCSFDITSTSTVDTKLSPFCDLTFHDEWIHYDLSQSLK 300
 QY 301 KYGYGAGNPLGPAAGGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGYGAGNPLGPAAGGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 QY 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420
 Db 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420
 QY 421 LVRVLNDRVPLHGGGVYDKLGRCKRDDEVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGGVYDKLGRCKRDDEVEGLSFARSGGNWEECEFA 467
 RESULT 2
 JN0656
 3-phytase (EC 3.1.3.8) A precursor - Aspergillus niger
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phya protein
 C:Species: Aspergillus niger
 C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
 A:Accession: JN0656; S28456
 R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.
 A.M.J.J.
 Gene 127, 87-94, 1993
 A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)
 A:Reference number: JN0656; MUID:93252284
 A:Accession: JN0656
 A:Molecule type: DNA
 A:Residues: 1-467 <NAN>
 A:Cross-references: GB:216414; MID:g2392; PIDN:CAA78904.1; PID:g2393
 A:Experimental source: strain NRRL3135
 A:Note: parts of the sequence, including the amino end of the mature protein, were confi
 C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and ino
 C:Genetics:
 A:Gene: phya
 A:Introns: 15/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-467/Product: 3-phytase A #status experimental <MAT>
 F:27-59,105,120,207,339,352,376,388/Binding site: carbohydurate (Asn) (covalent) #sta
 F:81,361/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 74.9%; Score 1848; DB 1; Length 467;
 Best Local Similarity 74.3%; Pred. No. 2.2e-137;
 Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLLSTATLFGSTGALCPNGNSHSCDVTGQYQCFPEISHLWGTYSPPFSLADE 60
 Db 1 MGVSALLLPLLYLLSGVTSLGLVAPASRNQSCDVTGQYQCFSETSHLWQYAPFSLANE 60
 QY 61 SATSPDVPKGRVTFVQVLSRGARYPTSSASKAYSALIEATQKNATAPKGYAFLKTYN 120
 Db 61 SVISPEVPAGCRVTFEAVLSRGARYPTDSKGYKYSALIEEQQNTTFDGYAFLKTYN 120
 QY 121 YTLGADLTPFGQPMVNSGIKRYKALARKIVPFRASGSDRVASAEKFIQGFOSA 180
 Db 121 YSLGADLTPFGQELVNSGIKRYQIESLTRNIVPFRSSSSRVASGKFKIEGFQST 180
 QY 181 KLADPGANPHQASPVINVIIEGAGYNNLTDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPRAQPCQSSPKIDVISEASSNNTLDPCTCTVFEDSELADTVEANFTAVPPIR 240
 QY 241 ARLEAHLPGVNLTDVNNLMDMCPDVTARSDATQLSPPFCDLTFHDEWIHYDLSQSLG 300
 Db 241 ORLENDLSGVTLTDEVTYLLMDMCSFDITSTSTVDTKLSPFCDLTFHDEWIHYDLSQSLK 300
 QY 301 KYGYGAGNPLGPAAGGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGYGAGNPLGPAAGGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 QY 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420
 Db 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420

QY 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420
 Db 361 HDNTMVSIFFFALGLYNGKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCAEKEP 420
 QY 421 LVRVLNDRVPLHGGGVYDKLGRCKRDDEVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGGVYDKLGRCKRDDEVEGLSFARSGGNWEECEFA 467
 RESULT 3
 JN0482
 3-phytase (EC 3.1.3.8) A - Aspergillus ficuum
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phya protein
 C:Species: Aspergillus ficuum
 C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
 C:Accession: JN0482; PN0023
 R:Ullah, A.H.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 192, 747-753, 1993
 A:Title: Aspergillus ficuum phytase: Complete primary structure elucidation by chemi
 A:Reference number: JN0482; MUID:93249451
 A:Accession: JN0482
 A:Molecule type: Protein
 A:Residues: 1-441 <ULL>
 A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
 R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 178, 45-53, 1991
 A:Title: Cyclohexanedione modification of arginine at the active site of Aspergillus
 A:Reference number: PN0023; MUID:91298982
 A:Accession: PN0023
 A:Molecule type: protein
 A:Residues: 48-70 <UL2>
 C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
 F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydurate (Asn) (covalent) #sta
 F:58,338/Active site: Arg, His #status predicted
 F:59/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 72.4%; Score 1787; DB 1; Length 441;
 Best Local Similarity 75.8%; Pred. No. 1.3e-132;
 Matches 332; Conservative 40; Mismatches 66; Indels 0; Gaps 0;

QY 27 NSHSCDVTGQYQCFPEISHLWGTYSPPFSLADESAISPDVPKGRVTFVQVLSRGARY 86
 Db 4 NQSSCDVTGQYQCFSETSHLWQYAPFSLANESVISPEVPAGCRVTFEAVLSRGARY 63
 QY 87 PTSASAKAYSALIEATQKNATAPKGYAFLKTYNLTGADLTPFGQPMVNSGIKRYR 146
 Db 64 PTDSKGRKYSALIEEQQNTTFDGYAFLKTYNLSGADLTPFGQELVNSGIKRYQ 123
 QY 147 YKALARKIVPFRASGSDRVASAEKFIQGFSAKLADPGANPHQASPVINVIIEGAGY 206
 Db 124 YESLTRNIVPFRSSSSRVASGKFKIEGFQSTKLDPRAQPCQSSPKIDVISEASS 183
 QY 207 NNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRLEAHLPGVNLTDVNNLMDMCP 266
 Db 184 NNTLDPGTCTVFEDSELADTVEANFTAVFPIRQRLNDELNSGVTLTDEVTYLLMDMCS 243
 QY 267 DTAVRTSDATQLSPPFCDLTFHDEWIHYDLSQSLGKYYGYGAGNPLGPAQCGVFNELIAR 326
 Db 244 DTISTSTVDTKLSPFCDLTFHDEWIHYDLSQSLKYYKYGAGNPLGPTQGVGVANELIAR 303
 QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFFALGLYNGKPLSTTSV 386
 Db 304 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFFALGLYNGKPLSTTSV 363
 QY 387 ESIEETDGYASWTVPFAARAYVEMMOCAEKEPFLVRLVNDRVVPLHGGGVYDKLGRCKR 446
 Db 364 ENITQDGFSSAWTVPFAARAYVEMMOCAEKEPFLVRLVNDRVVPLHGGGVYDKLGRCKR 423
 QY 447 DDFVEGLSFARSGGNWEE 464

A:Reference number: S53458
A:Accession: S53476
A:Molecule type: DNA
A:Residues: 1-467 <BUS>
A:Cross-references: EMBL:L28920; NID:g1616966; PIDN:AAC09508.1; PID:g456155; MIPS:YAR071
R:Chen, J.Y.; Gong, Y.; Ao, S.Z.
Acta Biochim. Biophys. Sin. 21, 437-444, 1989
A:Title: The primary structure of acid phosphatase gene PHO11 in *S. cerevisiae* and comparison with other yeast acid phosphatase genes
A:Reference number: JC1018
A:Accession: JC1018
A:Molecule type: DNA
A:Residues: 1-16, 'L', '18-149, 'H', '151-353, 'Q', '355-422, 'G', '424-467 <CHE>
A:Note: this paper is in Chinese, with an English abstract
C:Genetics:
A:Gene: SGD:PHO11
A:Cross-references: SGD:S0000094; MIPS:YAR071w
A:Map position: 1R
A:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-467/Product: acid phosphatase #status predicted <MAP>
F:74/Active site: Arg #status predicted
F:75/Active site: His (phosphohistidine intermediate) #status predicted
F:97,162,192,250,315,356,439,445,461/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.0%; Score 394; DB 2; Length 467;
Best Local Similarity 25.7%; Pred. No. 3.7e-23;
Matches 119; Conservative 65; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCPPEISHLWGTYSPPFSLADESAISDPVPGKGRVTFVQVLSRHG 83
DB 21 PLGKLSDDIKGTQTEIFPFL-----GGSGPYSPFGDYGISRDLPECEMKQVQVGRHG 76

QY 84 ARPTSSASKAYSAALIEAQKNATAPKGYAPLK-----TNYTL-----GADDL 128
DB 77 ERYPTVSKAKSIMTWTYKLSNVTGQFSGALSFNDDEFFIRDTKNLEMETTLANSVNVL 136

QY 129 TPF-GEQMVNSGIFRYRYKALARKIVPF-IRASGSDRVIASAEKFEIGFOSAKLADPG 186
DB 137 NPYTGEMNAKRHARDFLAQGYGVNQTSEFVTSNSNRCHDTAQYFDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFPESELGDDVEANFTAVFAPPIRA 241
DB 191 -----FNISLQPISEAESAGANTLSAHSCPAWDD-DVNDLKKYDKYLSGLIAK 240

QY 242 RLEAHLPGVNLTDVNNLMDCPDFTVARTSDATQLSPFCDLTHDEWIQYDYLSLQSGK 301
DB 241 RLKKNKGLNLTSSDANTFFAWCAYEINAR-----GYSDICNIFTKDELVRFSYGODLET 295

QY 302 YGYGAGNPLGPAQGVGVNELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSH 361
DB 296 YYQTGPGYDVRSVSGANLFNASVKLLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFGALNGTKPLSTTSVESIEETDGYSGASWTVPFAARAYVEMMOCEAEKEPL 421
DB 338 DTDILNLTITIGIIDDKNLTAHEVPMENT--FHRSWVYVPOGARVYTEKFQC--SNDTY 393

QY 422 VRVLVNDRVVPLHCGGVDKLGRCRDPD-----VEGLSFAR 457
DB 394 VRYVINDAVVPIETCTGPGFSCFCEINDFYAEKRVAGTDFLK 436

RESULT 6
S48996
acid phosphatase (EC 3.1.3.2) PHO12 - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YHR215w
C:Species: *Saccharomyces cerevisiae*
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999
C:Accession: S48996; S59659
R:Macri, C.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9177.

A:Reference number: S46671
A:Accession: S48996
A:Molecule type: DNA
A:Residues: 1-467 <MAC>
A:Cross-references: EMBL:U00029; NID:g551322; PIDN:AAB69729.1; PID:g458917; MIPS:YHR2
R:Yu, L.
submitted to the EMBL Data Library, January 1995
A:Reference number: S59658
A:Accession: S59659
A:Molecule type: DNA
A:Residues: 1-16, 'L', '18-81, 'AR', '84-149, 'H', '151-467 <XUL>
A:Cross-references: EMBL:U19789; NID:g847754; PIDN:AAA73479.1; PID:g847755
C:Genetics:
A:Gene: SGD:PHO12
A:Cross-references: SGD:S0001258; MIPS:YHR215w
A:Map position: 8R
A:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 16.0%; Score 394; DB 2; Length 467;
Best Local Similarity 25.7%; Pred. No. 3.7e-23;
Matches 119; Conservative 65; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCPPEISHLWGTYSPPFSLADESAISDPVPGKGRVTFVQVLSRHG 83
DB 21 PLGKLSDDIKGTQTEIFPFL-----GGSGPYSPFGDYGISRDLPECEMKQVQVGRHG 76

QY 84 ARPTSSASKAYSAALIEAQKNATAPKGYAPLK-----TNYTL-----GADDL 128
DB 77 ERYPTVSKAKSIMTWTYKLSNVTGQFSGALSFNDDEFFIRDTKNLEMETTLANSVNVL 136

QY 129 TPF-GEQMVNSGIFRYRYKALARKIVPF-IRASGSDRVIASAEKFEIGFOSAKLADPG 186
DB 137 NPYTGEMNAKRHARDFLAQGYGVNQTSEFVTSNSNRCHDTAQYFDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFPESELGDDVEANFTAVFAPPIRA 241
DB 191 -----FNISLQPISEAESAGANTLSAHSCPAWDD-DVNDLKKYDKYLSGLIAK 240

QY 242 RLEAHLPGVNLTDVNNLMDCPDFTVARTSDATQLSPFCDLTHDEWIQYDYLSLQSGK 301
DB 241 RLKKNKGLNLTSSDANTFFAWCAYEINAR-----GYSDICNIFTKDELVRFSYGODLET 295

QY 302 YGYGAGNPLGPAQGVGVNELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSH 361
DB 296 YYQTGPGYDVRSVSGANLFNASVKLLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFGALNGTKPLSTTSVESIEETDGYSGASWTVPFAARAYVEMMOCEAEKEPL 421
DB 338 DTDILNLTITIGIIDDKNLTAHEVPMENT--FHRSWVYVPOGARVYTEKFQC--SNDTY 393

QY 422 VRVLVNDRVVPLHCGGVDKLGRCRDPD-----VEGLSFAR 457
DB 394 VRYVINDAVVPIETCTGPGFSCFCEINDFYAEKRVAGTDFLK 436

RESULT 7
PABYCC
acid phosphatase (EC 3.1.3.2) precursor, constitutive - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c
C:Species: *Saccharomyces cerevisiae*
C:Date: 30-Sep-1991 #sequence_revision 09-Sep-1994 #text_change 05-Nov-1999
C:Accession: S48259; S05794; A25241; S44674
R:Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.
Yeast 10, 1363-1381, 1994
A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A:Reference number: S48255; MUID:95208357
A:Accession: S48259
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-467 <MAN>
A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55597.1; PID:g476050

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
 R:Feidmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
 Submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45927

A:Molecule type: DNA
 A:Residues: 1-467 <PE2>
 A:Cross-references: EMBL:Z35961; NID:g536362; PIDN:CAA85045.1; PID:g536363; GSPDB:GN0000
 R:Bajwa, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
 Nucleic Acids Res. 12, 7721-7739, 1984
 A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeast
 A:Reference number: S05794; MUID:85037940

A:Accession: S05794
 A:Molecule type: DNA
 A:Residues: 1-218, 'MKT', 222-467 <BAJ1>
 A:Cross-references: EMBL:X01080; NID:g4148; PIDN:CAA25557.1; PID:g758281
 A:Note: the authors translated the codon AAT for residue 134 as Asp and TAC for residue
 R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G.
 Mol. Cell. Biol. 6, 1855-1865, 1986
 A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
 A:Reference number: A93074; MUID:87064474
 A:Accession: A25241
 A:Molecule type: DNA
 A:Residues: 1-44 <TAI>

C:Genetics:
 A:Gene: SGD:PHO3; MIPS:YBR092c
 A:Cross-references: SGD:S0000296; MIPS:YBR092c
 A:Map position: 2R

C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-467/Product: acid phosphatase, constitutive #status predicted <MA>
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (covall
 F:337/Active site: His #status predicted

Query Match 15.8%; Score 391; DB 1; Length 467;
 Best Local Similarity 27.0%; Pred. No. 6.3e-23;
 Matches 117; Conservative 61; Mismatches 192; Indels 64; Gaps 13;

QY 49 GTYSPFSLADESAISPDVPGKRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATA 108
 DB 42 GGAGYFSPGPDYGISRLPECEMKQQLARHGERPTYSKGTATMKTWYKLSNTRQ 101
 QY 109 FKGKVAFLK-TYNYTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 DB 102 FNGSLFNDYEFFIRDDDDLEMETTFANSNDVNLNPTGEMDAKRAHAREFLAQGYMFE 161
 QY 153 KIVPP-IRASGSDRVIAAEKFIQFQSAKLADPGANPHQASPVINVIIEGAGYNTLD 211
 DB 162 NOTSPPIFAASERVHDTAQYFDGL-----GDQFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFESEELGDVDEANFTAVFAPP-IRARLAHLPGVNLTDDEVVNLMDMCPPTVA 270
 DB 211 AGNACPGWDE-DANDDILDKYDVTYLLDIAKRLNKENGLNLTSDKANTLFWCAYELNA 269
 QY 271 RTSDATQLSPCLDTHDEWTDYDYLQSLGKYGYGAGNPLGPAQGVGVNELIARLTHS 330
 DB 270 R-----GYSDVCDLTFEDELRYVSGDGLVSFYODGPGYDMIRSVGANLENATLKLKQS 324
 QY 331 PVQDHTSTNHTLDSNPATFLNATLYADFSDHNTMVSIFALGLYNGTKPLSTVSIESIE 390
 DB 325 ETQD-----LKWLSFTHTDILNLTAGLIDKNNLTAEYVPPFMG 366
 QY 391 ETDGYSASWTVPFAARAYVEMMOCAEKEPLVRVLVNDRVVPLHGGCVDKLGRCKRDFE- 449
 DB 367 NT--FHKSWSYVPOGARVYTEKFQC--SNDYVVRVINDAVVPIETCTGTGFCSEINDFY 422
 QY 450 -----VEGLSFAR 457

DB 423 DYAEKRVAGTDFLK 436

RESULT 8

S52495
 acid phosphatase homolog YDL024c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D2815
 C:Species: Saccharomyces cerevisiae

C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
 C:Accession: S52495; S67556
 R:Andre, B.; Vissers, S.; Urrestazu, L.

submitted to the EMBL Data Library, February 1995
 A:Description: The sequence of a 42 kb segment located on the left arm of chromosome

A:Reference number: S52492

A:Accession: S52495

A:Molecule type: DNA

A:Residues: 1-468 <AND>

A:Cross-references: EMBL:Z48432; NID:g683669; PIDN:CAA88335.1; PID:g683673

R:Urrestazu, L.A.; Andre, B.; Vissers, S.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67535

A:Accession: S67556

A:Molecule type: DNA

A:Residues: 1-468 <URR>

A:Cross-references: EMBL:Z74072; NID:g1430996; PIDN:CAA98583.1; PID:g1430997; MIPS:YD
 A:Experimental source: strain S288C

C:Genetics:

A:Map position: 4L

C:Superfamily: yeast acid phosphatase

Query Match 15.0%; Score 369; DB 2; Length 468;
 Best Local Similarity 25.3%; Pred. No. 3.4e-21;
 Matches 117; Conservative 60; Mismatches 175; Indels 110; Gaps 14;

QY 49 GTYSPFSLADESAISPDVPGKRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATA 108

DB 43 GGAGYFSPFANYGIPDTIPEGCRLTQVQMGIRGERTSEAKDIFEVWYKISNVTGK 102

QY 109 FKGKVAFLKT-YNYTL-----GADDLTPF-GEQMVNSGKIFRYRYKALAR 152

DB 103 YEGSLFLNNGYEFFIPDESLEMETTLQNSIDVLPYTGEMNAKRAHAREFLAKYKGLME 162

QY 153 KIVPP-IRASGSDRVIAAEKFIQFQSAKLADPGANPHQASPVINVIIEGAGYNTLD 211

DB 163 NCTNFPIFTTNSKRIYDTAQYFAEL-----GDQFNISLQ 197

QY 212 HGLCTAFESEELG-----DDVEANFTAVFAPPTRARLEAHLPGVNLT 253

DB 198 -----TLSESSSSGANTLAAKSSCPNWNANNDILMSYSDYLENIDRLNDENKGLNS 253

QY 254 DEDVNLMDMCPPTIVARTSDATQLSPFCDLFTHDEWTDYDYLQSLGKYGYGAGNPLGP 313

DB 254 RKDAALFSWCAEL-----NAKGYSNICDIFSAELIHYSETDILTSTFYQNGPGYKLIK 308

QY 314 AQGVGVN---ELIARLTHSPVQDHTSTNHTLDSNPATFLNATLYADFSDHNTMVSIF 370

DB 309 SIGANLFNATVKLIROSAH-----LDQKWLSTHTDIDILNYLT 347

QY 371 ALGLYNGTKPLSTVSIESIEEDGYASWTVPFAARAYVEMMOCAEKEPLVRVLVNDRV 430

DB 348 TAGLIDDTNLTNNHVPFRDHS--YHRSWYIPQGARVYTEKFQC--SNDYSYRVVNDVAV 403

QY 431 VPLHGGCVDKLGRCKRDFE-----GLSFARSGGNWEEC 465

DB 404 VPIESSCGPGSCERGTFYAKDLRGVSP-----YEDC 439

RESULT 9

JN0715

3-phytase (EC 3.1.3.8) B precursor - Aspergillus ficum

N:Alternate names: pH 2.5-optimum acid phosphatase

C:Species: Aspergillus ficum

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

[illegible]

a:Gene: phoI; SPDB:SPBP4G3.02
 a:Map position: 2
 a:Cross-references: EMBL:AL137099; PIDN:CAB68657.1; GSPDB:GN00067; SPDB:SPBP4G3.02
 a:Experimental source: strain 972h(-); clone p1 p4G3
 a:Genetics:

```

89 LNNFQEKLLNGSIPNVFSYPENPLCFIKQWTPVIDAENADQLSSRRGLEFDLGRQLYQR 148
147 YKALARKTVPFETRAGSSDRVTASAEKTEGEOCAKI ADPCANDHOACSDVTNVTIDEC 204

```

Db 149 YKLFDSYVDINTAEQKRVESAKWTFYGLFGDKWYE-----KTNFILLISEGKAA 199
 QY 205 GYNNTLDHGLCTAFBESEL----GDDVEANETAVFAPPPIRARLEAHL-PGVNLTDEDVYN 259
 Db 200 GANSLSMYNACPVFKDNNPHKKNATDAHAHVNRNFIETPVNRLAKYFDSYKLTINDVRS 259
 QY 260 LDMCPCFDVARTSDATQSPFCDLFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGF 319
 Db 260 LEVICEYEIAIKDH-----SDFCSLFTPSEFLNFYDSDLDQAYGGGPVSEWASTLGGAY 314
 QY 320 VNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTK 379
 Db 315 INNLADSL-----RNVTNPFDRK-----VFLAFTHDSNLIIPVEAALGFPPDIT 358
 QY 380 PLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAEKEPLVRLVNDVRVPLHGCCVD 439
 Db 359 PQNPLPTDKNIITYSOKTSSFPFAGNLITELFFC-SDSKYYVRLVNOQVYPLIDCGYG 417
 QY 440 KLGR-----CKRDDFVGLSFARSGGN 461
 Db 418 PSGTSDGLCELQAYLNSPIRANSTN 443

RESULT 14

SL1419
 A:Title: acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
 C:Accession: SL1419; T40455
 R:Yang, J.; Schweingruber, M.E.
 Curr. Genet. 18, 269-272, 1990
 A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schizosaccharomyces pombe
 A:Reference number: SL1419; MUID:91064763
 A:Accession: SL1419

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <YAN>
 A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA0258.1; PID:g5007
 R:lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z21931
 A:Accession: T40455
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-463 <LYN>
 A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN00067; SPDB:SPBC428.03c
 A:Experimental source: strain 972h; cosmid c428

C:Genetics:
 A:Gene: SPBC428.03c
 A:Map position: 2
 A:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 12.3%; Score 303.5; DB 2; Length 463;
 Best Local Similarity 26.7%; Pred. No. 4.7e-16;
 Matches 113; Conservative 63; Mismatches 184; Indels 63; Gaps 16;

QY 46 HLMGTYS-----PFFSLADESAISPDVPKGRVTFVQVLSRHGARYPT--SSASKAYSAA-- 97
 Db 35 HL-GTISVYHEPYF-----NGPTTSFPESCAIKQVHLLQHRGSRNPTGDDTATDVSSAQY 88
 QY 98 -----LIEATOKNATAFKYAFKTYNTL-----GADDLTPFGEQOOWNSGKIFRR 146
 Db 89 IDIFQNLKNGSIPVNFSPENFLYFVKHTVPVKEANDQISSLSSRIELFDLGRGVFR 148
 QY 147 YKALARKIVPFIIRASGSDRVIAAEKFIIEGFSQAKLADPAPNPHQASPVNVIPIE--GA 204
 Db 149 YYELFDFTDVIDYNTAAQERVVDSEAFNSYGM-----FGDDMQNKTNF-----IVLPEDDSA 199
 QY 205 GYNNTLDHGLCTAFBESELGDD--VEANFTA---VFAPPPIRARLEAHL-PGVNLTDEDVYN 259

Db 200 GANSLAMYSCPVEYEDNNIDNTTEAAHTSWRNVLFLKPIANRLNKYFDSGYNLTVDVRS 259
 QY 260 LDMCPCFDVARTSDATQSPFCDLFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGF 319
 Db 260 LYVICVYEIALRDN-----SDFCSLFTPSEFLNFYDSDLDQAYGGGPVSEWASTLGGAY 314
 QY 320 VNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTK 379
 Db 315 VNNLANNL-----RKGYNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDIT 358
 QY 380 PLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAEKEPLVRLVNDVRVPLHGCCVD 439
 Db 359 PEHPLPTDKNIITYSOKTSSFPFAGNLITELFFC-SDSKYYVRLVNOQVYPLIDCGYG 417
 QY 440 KLGR 442
 Db 418 PSG 420

RESULT 15

JE0369
 A:Title: histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
 C:Accession: JE0369
 R:Mullaney, E.J.; Ullah, A.H.J.
 Biochem. Biophys. Res. Commun. 251, 252-255, 1998
 A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidopsis

A:Reference number: JE0369; MUID:99009256
 A:Accession: JE0369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-465 <MUL>
 A:Superfamily: Yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 6.2%; Score 153; DB 2; Length 465;
 Best Local Similarity 22.5%; Pred. No. 0.00031;
 Matches 102; Conservative 61; Mismatches 199; Indels 92; Gaps 21;

QY 35 DGYQCFPEITSHLWGT---YSPFSLADESAISPDVPKGRVTFVQVLSRHGARYPTSSA 91
 Db 16 DGGF---DHRHLSTVTRYSTSKDVTQNLIEGSRNPSECTPIHLNLVLRHGRSTRSPTKR 71
 QY 92 SKAYSAL---LEATOKNATAFK-----GKYAFKTYNTLTGADDLTPFGEQOOWN 138
 Db 72 LREMESLAGRFKELVROAEARKLPDSKIPGWLQWK--SPWEGKVKGELIROGEDELYQ 129
 QY 139 SGIKFYRRYKALARK---IVPFIIRASGSDRVIASAEKFIIEGFSQAK-LADPGANPHQAS 193
 Db 130 LGIRVREFPFLFEEDYHPDVYTTIRATQIPRASASAVAFGMLFSEKGNLGPGRNRAFA- 188
 QY 194 PVINVIPEGAGYNTLHDGLCTAFBESELGDDVEANFTAVFAPPPIRARLEAHLPGVNLT 253
 Db 189 -----VTSNRASDTKURFFEC-----QNV-----KSYRAKEPAVDKL 223
 QY 254 DEDVNLMDMC---PFDVARTSDATQSPFCDLFTHDEWIOYDYLQSLGKYGYGAGNP 310
 Db 224 KEPVNLKITASVAKRYDLKFTKQDISSLWFLCKQVALLW-----TDDLEVLKKGYSN 278
 QY 311 LGPAGGVFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMWSIFF 370
 Db 279 LNYKMGVPLLEDVL---HSMEEAIKAREEKLK--PGSYE-KARL--RFAHAETIVPPFS 329
 QY 371 A-LGLYNTKPLSTTSVESIE-----ETDGYASWTVPFAARAYVEMMOCAE 417
 Db 330 CULGLF-----LDGEFEKIQKEKPELPPOPKTRFRGCTMAPFGNNLLVLYSCP 384
 QY 418 KEP--LVRLVNDVRVPLHGCCVDKLGKCRDDF 449
 Db 385 SSPKYFVQVHLNHEHPIAVPGC--DGKDFCPLDEF 416

Sat Oct 27 15:25:55 2001

us-09-488-265-31.rpr

Page 9

Search completed: October 26, 2001, 15:10:30
Job time: 6336 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:18:04 ; Search time 29.88 Seconds
(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-31

Perfect score: 2468.

Sequence: 1 MGFFVLLSTALFGTSGT.....DFVEGLSFARSGGNWERCFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	1852	75.0	467	PHYA_ASPAW	P34753 aspergillus
2	1848	74.9	467	PHYA_ASPNG	P34752 aspergillus
3	1804.5	73.1	463	PHYB_EMENI	O00093 emericeella
4	405	16.4	467	PHY5_YEAST	P00635 saccharomyc
5	394	16.0	467	PPAB_YEAST	P35842 saccharomyc
6	394	16.0	467	PPAC_YEAST	P38693 saccharomyc
7	391	15.8	467	PPA3_YEAST	P24031 saccharomyc
8	369	15.0	468	PPAD_YEAST	P52290 saccharomyc
9	362.5	14.7	479	PHYB_ASPNG	P34754 aspergillus
10	354.5	14.4	479	PHYB_ASPAW	P34755 aspergillus
11	350	14.2	468	PPAL_PICPA	P52291 pichia past
12	337.5	13.7	469	PPA5_KLULA	P52289 kluyveromyc
13	328	13.3	453	PPA1_SCHPO	P08091 schizosacch
14	303.5	12.3	463	PPA2_SCHPO	Q01682 schizosacch
15	142	5.8	755	PPAX_CAEEL	Q09549 caenorhabdi
16	127.5	5.2	423	PPAL_RAT	P20611 rattus norv
17	118.5	4.8	421	PPAL_MOUSE	P24638 mus musculus
18	115.5	4.7	381	PPAP_RAT	P20646 rattus norv
19	114.5	4.6	423	PPAL_HUMAN	P11117 homo sapien
20	106.5	4.3	602	PEHX_ERWCH	P15922 erwinia chr
21	101.5	4.1	413	AGP_ECOLI	P19926 escherichia
22	99	4.0	386	PPAP_HUMAN	P15309 homo sapien
23	98	4.0	844	YD25_SCHPO	O13712 schizosacch
24	97.5	4.0	693	LYS4_YEAST	P43367 saccharomyc
25	97	3.9	411	PPAY_CAEEL	Q10944 caenorhabdi
26	96.5	3.9	845	CC47_YEAST	P38132 saccharomyc
27	96	3.9	721	NTCA_CAEEL	Q23316 caenorhabdi
28	95.5	3.9	609	COPA_PESM	P12374 pseudomonas
29	94.5	3.8	526	CLOS_CLOHI	P09870 clostridium
30	94	3.8	467	APPX_BACFI	P25075 bacillus fi
31	94	3.8	938	TRGI_ECOLI	P33790 escherichia
32	93.5	3.8	2485	PTND_HUMAN	O12923 homo sapien
33	93	3.8	428	GSA_STAAU	O34092 staphylococ

34 93 777 1 ISOA_FLASP
35 92.5 433 1 DHOM_VSNY3
36 92.5 3.7 760 1 MCM7_SCHPO
37 92 3.7 3587 1 TYCB_BACBR
38 91.5 3.7 351 1 PYRD_HELPY
39 91.5 3.7 586 1 VP40_ILTVT
40 91.5 3.7 680 1 NCPR_CANMA
41 91 3.7 417 1 AGP_PRORE
42 90.5 3.7 646 1 IRGI_MOUSE
43 90 3.6 626 1 BGAL_LEULA
44 90 3.6 1527 1 MRP3_HUMAN
45 89.5 3.6 492 1 CRTI_ERWHE

ALIGNMENTS

RESULT 1
ID PHYA_ASPAW STANDARD; PRT; 467 AA.
AC P34753:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA OR PHY.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALK0243;
RX MEDLINE=94040796; PubMed=8224894;
RA Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
RA Miettinen-Oinonen A., Nevalainen H., Rambossek J.A.;
RT "The cloning and sequencing of the genes encoding phytase (phy) and
RT pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
RT awamori.";
RL Gene 133:55-62(1993).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H2O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L02421; AAA16898.1;
CC PIR; JN0889; JN0889.
CC HSP; P34752; IHP.
CC InterPro; IPR000560;
CC Pfam; PF00328; acid.phosphat; 1.
CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC Hydrolyase; Glycoprotein; Signal.
CC SIGNAL 1 23 POTENTIAL A.
CC CHAIN 24 467 3-PHYTASE A.
CC FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
CC FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
CC FT DISULFID 31 40 BY SIMILARITY.
CC FT DISULFID 71 414 BY SIMILARITY.
CC FT DISULFID 215 465 BY SIMILARITY.
CC FT DISULFID 264 282 BY SIMILARITY.
CC FT DISULFID 436 444 BY SIMILARITY.

```
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 51075 MW; 118E828A5D7EC661 CRC64;

Query Match 75.08; Score 1852; DB 1; Length 467;
Best Local Similarity 74.38; Pred. No. 2.1e-140;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGCVFVLLSIATLGSSTGALGPRGNHSCDVTGQYQCFPEISHLWGYSPFFSLADE 60
DB 1 MGVSALLPLLYLAGVTSGLVAPASRNQSTCDTVDQGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAFKGKVAFLKTYN 120
DB 61 SAISPDVPAGCRVTFQAQVLSRHGARYPTESKGYKYSALIEIQQNVTFDGKVAFLKTYN 120

QY 121 YTLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPIFRASGDRVIAAEKFTIEGFOSA 180
DB 121 YSLGADLLTPFGEQELVNSGKIFRYRYKALARKIVPIFRASGDRVIAAEKFTIEGFOST 180

QY 181 KLADPGANPHQASPVNIIIEGAGYNNTLDHGLCTAFESSELGDDVEANFTAVFAPPIR 240
DB 181 KLKDPRAQGGSSPKIDVIVSEASSNNTLDPGCTVFESSELADTVEANFTAVFAPSI 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDTVARTSDATLSPDCDLTFDHEMIOXDYDLSLG 300
DB 241 QLENDLSGVLTDEVTYLMDCSFDTISTSTVDTKLSFPCDLTFDHEMIDHYDLSLK 300

QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNNHLDSPNATFPLNATFYADFS 360
DB 301 KYGHCAGNPLGTPQGVYANELIARLTHSPVHDTSSNHLDSNPATFPLNSTLYADFS 360

QY 361 HDNTMVSITFFALGNYGKPLSTTSVESIEETDGYSAWVTFPAARAYVENMQCEAKEP 420
DB 361 HDNGIISILFALGNYGKPLSTTTVENITDGFESSAWTVPFASRLYVENMQCAEQEP 420

QY 421 LVRVLNDRVPLHCGVDKLGCRKRDDEFVGLSFAKSGGNWBCFA 467
DB 421 LVRVLNDRVPLHCGPIDALGCRTRDSFVGLSFAKSGGDWAECSA 467

RESULT 2
ID PHVA_ASPNG STANDARD; PRT; 467 AA.
AC P34752;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHVA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93252284; PubMed=8387447;
RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
RA Gouka R.J., Suykerbuyk M.E.G., Luiten R.G.M., van Paridon P.A.,
RA Selden G.C.M., Veenstra A.E., van Gorcom R.F.M.,
RA van den Hondel C.A.M.J.J.;
```

```
RT *Cloning, characterization and overexpression of the phytase-encoding
RT gene (phyA) of Aspergillus niger.";
RL Gene 127:87-94(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mullaney E.J.;
RL Submitted (XXI-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-464.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93249451; PubMed=8387289;
RA Ullah A.H.J., Dischinger H.C. Jr.;
RT "Aspergillus ficum phytase: complete primary structure elucidation
RT by chemical sequencing.";
RL Biochem. Biophys. Res. Commun. 192:747-753(1993).
RN [4]
RP SEQUENCE OF 71-93.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=91298982; PubMed=1648914;
RA Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
RT "Cyclohexanediol modification of arginine at the active site of
RT Aspergillus ficum phytase.";
RL Biochem. Biophys. Res. Commun. 178:45-53(1991).
RN [5]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=89160685; PubMed=2852807;
RA Ullah A.H.J.;
RT "Aspergillus ficum phytase: partial primary structure, substrate
RT selectivity, and kinetic characterization.";
RL Prep. Biochem. 18:459-471(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=97307250; PubMed=9164457;
RA Kostrewa D., Gruning-Leitch F., D'Arcy A., Broger C., Mitchell D.,
RA van Loon A.P.;
RT "Crystal structure of phytase from Aspergillus ficum at 2.5-A
RT resolution.";
RL Nat. Struct. Biol. 4:185-190(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN
CC FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
CC USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER
CC THE NAME PHYTASE NOVO.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z16414; CAA78904.1; -
DR EMBL; M94550; AAA32705.1; -
DR PIR; JN0482; JN0482.
DR PIR; JN0023; JN0023.
DR PIR; JN0656; JN0656.
DR PDB; 1IHP; 18-MAR-98.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 467 3-PHYTASE A.
FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
```

```

FT DISULFID 31 40
FT DISULFID 71 414
FT DISULFID 215 465
FT DISULFID 264 282
FT DISULFID 436 444
FT CARBOHYD 27 27
FT CARBOHYD 59 59
FT CARBOHYD 105 105
FT CARBOHYD 120 120
FT CARBOHYD 207 207
FT CARBOHYD 230 230
FT CARBOHYD 339 339
FT CARBOHYD 352 352
FT CARBOHYD 376 376
FT CARBOHYD 388 388
SQ SEQUENCE 467 AA; 51086 MW; 88FE8F3584341D6D CRC64;

Query Match 74.9%; Score 1848; DB 1; Length 467;
Best Local Similarity 74.3%; Pred. No. 4.5e-140;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSLGLAVPASRQSSCDTVGQYCFSETHLWGYAPFFSLANE 60
Db 1 MGVSALLPLLYLLSGVTSLGLAVPASRQSSCDTVGQYCFSETHLWGYAPFFSLANE 60

QY 61 SATSPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFLKTYN 120
Db 61 SATSPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQOMVNSGKFKYRRYKALARKIVPFRASGSDRVIASAEKFEQFQSA 180
Db 121 YTLGADDLTPFGQOMVNSGKFKYRRYKALARKIVPFRASGSDRVIASAEKFEQFQSA 180

QY 181 KLADPGANPHQASPVINVIIPGAGYNNLTDLGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVINVIIPGAGYNNLTDLGLCTAFEESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLDEVDVNLDMCPDPTVARTSDATQSPFCDLFTHDEWIQDYLOSLG 300
Db 241 ARLEAHLPGVNLDEVDVNLDMCPDPTVARTSDATQSPFCDLFTHDEWIQDYLOSLG 300

QY 301 KYGAGNPLGPAQGVGFVNNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYGAGNPLGPAQGVGFVNNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMVSIFPFGALYNGTKPLSTTSVESIEETDGYASWTVTPFAARAYVEMMOCEAKEP 420
Db 361 HDNTMVSIFPFGALYNGTKPLSTTSVESIEETDGYASWTVTPFAARAYVEMMOCEAKEP 420

QY 421 LVRVLNDRVPLRGCGVDKLGRCRDDFVEGLSFARSGCNWEECA 467
Db 421 LVRVLNDRVPLRGCGVDKLGRCRDDFVEGLSFARSGCNWEECA 467

```

```

RESULT 3
ID PHYB_EMENI
AC 000033; STANDARD; PRT; 463 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHODIOLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHODIOLASE B).
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98007872; PubMed=9349716;
RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,

```

```

RA van Loon A.P.;
RT "Cloning of the phytases from Emericella nidulans and the
RL thermophilic fungus Talaromyces thermophilus.";
RL Biochim. Biophys. Acta 1353:217-223(1997).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U59803; AAB96871.1; -
CC InterPro: IPR000560; -
CC Pfam: PF00328; acid_phosphat; 1.
CC PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC KW Hydrolyase; Glycoprotein; Signal.
CC FT SIGNAL 1 19 POTENTIAL.
CC FT CHAIN 20 463 3-PHYTASE B.
CC FT ACT_SITE 80 80 REQUIRED FOR BINDING SUBSTRATE
CC FT ACT_SITE 81 81 (BY SIMILARITY).
CC FT ACT_SITE 357 357 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
CC FT CARBOHYD 26 26 PROTON DONOR (BY SIMILARITY).
CC FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 463 AA; 51786 MW; ECC5827D1E1C82A2 CRC64;

Query Match 73.1%; Score 1804.5; DB 1; Length 463;
Best Local Similarity 72.5%; Pred. No. 1.3e-136;
Matches 338; Conservative 48; Mismatches 75; Indels 5; Gaps 3;

QY 1 MGVSALLPLLYLLSGVTSLGLAVPASRQSSCDTVGQYCFSETHLWGYAPFFSLANE 60
Db 1 MAFFTVALSLLYLLSRVSAQA--PVVQNHSCNTADGGYQCFPNVSHVWGQYSPYFSIEQE 58

QY 61 SATSPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFLKTYN 120
Db 59 SATSEDPVHGCEVTFFQVLSRHGARYPTSSASKAYSALIEAIOKNATSFQVAFLESYN 118

QY 121 YTLGADDLTPFGQOMVNSGKFKYRRYKALARKIVPFRASGSDRVIASAEKFEQFQSA 180
Db 119 YTLGADDLTPFGQOMVNSGKFKYRRYKALARKIVPFRASGSDRVIASAEKFEQFQSA 178

QY 181 KLADPGANPHQASPVINVIIPGAGYNNLTDLGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 179 QLDHDSK--RATPVVNVIIPEIDGFNNTLDHSCVSEFENDERADETEANFTAIMGPPIR 236

QY 241 ARLEAHLPGVNLDEVDVNLDMCPDPTVARTSDATQSPFCDLFTHDEWIQDYLOSLG 300
Db 237 KRLENDLPGLKLTNENVIYLMDCSFDTMARTAGTSLSPFCALFTTEKEWLQDYLOSL 296

QY 301 KYGAGNPLGPAQGVGFVNNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
Db 297 KYGAGNPLGPAQGVGFVNNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 356

QY 361 HDNTMVSIFPFGALYNGTKPLSTTSVESIEETDGYASWTVTPFAARAYVEMMOCEAKEP 420
Db 357 HDNSMISIFPFGALYNGTKPLSTTSVESIEETDGYASWTVTPFAARAYVEMMOCEAKEP 415

```


ID AC P35842; STANDARD; PRT; 467 AA.
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE ACID PHOSPHATASE PHO11 PRECURSOR (EC 3.1.3.2) (P56).
 GN PHO11 OR YAR071W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=95249563; PubMed=7731988;
 RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
 RA Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
 RA Storms R.K.;
 RT "The nucleotide sequence of chromosome I from Saccharomyces
 cerevisiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen J.Y., Gong Y.I., Ao S.Z.;
 RT "The primary structure of acid phosphatase gene PHO11 in S.
 cerevisiae and comparison with other gene families.";
 RL Acta Biochim. Biophys. Sin. 21:437-444(1989).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
 CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L28920; AAC09508.1; -
 DR EMBL; U19789; AAA73479.1; -
 DR HSSP; P34752; 1IHP.
 DR SGD; S0000094; PHO11.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 17
 FT ACT_SITE 18
 FT ACT_SITE 75
 FT ACT_SITE 337
 FT CARBOHYD 97
 FT CARBOHYD 162
 FT CARBOHYD 192
 FT CARBOHYD 250
 FT CARBOHYD 315
 FT CARBOHYD 356
 FT CARBOHYD 390
 FT CARBOHYD 439
 FT CARBOHYD 445
 FT CARBOHYD 461
 FT CONFLICT 17
 FT CONFLICT 82
 FT CONFLICT 150
 FT CONFLICT 354
 FT CONFLICT 423
 FT CONFLICT 424
 FT CONFLICT 467 AA; 52757 MW; AECDIC046B326C3; CRC64;
 SQ SEQUENCE

Query Match 16.0%; Score 394; DB 1; Length 467;
 Best Local Similarity 25.7%; Pred. No. 5.4e-24;
 Matches 119; Conservative 65; Mismatches 203; Indels 76; Gaps 14;
 QY 24 PRGNHSCDVTGGYQCFPEISHLWGTYSPEFSLADESAISPDVPGKCRVTFVQVLSRHG 83
 DB 21 PLGKSLDIDKIGTQTEIFPEP-----GSGGYTSFPGYGISRLDPESCEMKQVMVGRHG 76
 QY 84 ARYPTSSAKSAYSALIEAIQKNATAFKGYAFLK-----TNYNLT-----GADDL 128
 DB 77 ERYPTSVSKAKSIMTWTYKLSNYTGQSGALSFLNDDVEFFIRDTKMLEMETTLANSVNL 136
 QY 129 TPF-GEQOMVNSGIKFYRRYKALKARIVPP-IRASSDRVIAEAEKFIQFQSAKLADPG 186
 DB 137 NPYTGEMNAKRHARDLQAQGYMVENQTSFAVFTSNRNCHRDTAQYFDG-----LGDK- 190
 QY 187 ANPHQASPVINVII-----PEGAGYNNLTLDHGLCTAFEESESLGDGDEANFTAVFAPPIRA 241
 DB 191 -----FNISLOTISEAESAGANTLSAHSCPAWDD-DVNDLILKKYDTKYLGIK 240
 QY 242 RLEAHLPGVNLDTEDVNNLMDMCPFDVTARTSDATQSPFCDLTHDEWIQYDYLQSLGK 301
 DB 241 RLKKNKGLNLTSSDANTFFAWCAYEINAR-----GYSDICNIFTKDELVRFSYGQDLET 295
 QY 302 YGYGAGNPLGPAQGYGVFNELIARLTHSPVQDHTSTNHTLDSNPATFFPLNATLYADFSH 361
 DB 296 YQYTGCGYDVVRSGANLFNASVKLKESEVQDQ-----KWLVSFTH 337
 QY 362 DNTMVSIFFALGNGTKPLSTTSVESIETDGYSTVTPFAARAYVEMMOCEAEKEPL 421
 DB 338 DTDILNLTIGIIDKNNLTAHEVPFMENT--FHSWYVPGQARVYTERKFCQ--SNDTY 393
 QY 422 VRVLYNDRVVPPLHGCGVDKLGCRKRDFF-----VEGLSFAR 457
 DB 394 VRVINDAVVPIETCTGPGFCEINFDYDAEKRVAGTDFLK 436
 RESULT 6
 PPAC_YEAST STANDARD; PRT; 467 AA.
 ID PPAC_YEAST
 AC P38693;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE ACID PHOSPHATASE PHO12 PRECURSOR (EC 3.1.3.2).
 GN PHO12 OR YHR215W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nham M., Rifkin L., Riley L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII.";
 RL Science 265:2077-2082(1994).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
 CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; U00029; AAB69729.1; -
 CC PIR; S48996; S48996.
 CC HSP; P34752; 1IHP.
 CC SCD; S0001258; PHO12.
 CC InterPro; IPR000560; -
 CC Pfam; PF00328; acid_phosphat; 1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 461 461
 FT SEQUENCE 467 AA; 52699 MW; BEC606CDF39B845B CRC64;

Query Match 16.0%; Score 394; DB 1; Length 467;

Best Local Similarity 25.7%; Pred. No. 5.4e-24; Mismatches 119; Conservative 65; Indels 76; Gaps 14;

QY 24 PRGNHSCDTPGVGQCPEIHLWGTVSPFSLADESAISPDPVKGRVTFVQVLSRHG 83
 DB 21 PLGKLSIDIKGTQTEIPFL---GGSGPYYSFFGDYGISRLPESCCKQVQVGRHG 76
 QY 84 ARPTSSASKAYSAIEAIOKNATFKYAPLK-----TNYTL-----GADDL 128
 DB 77 ERYPTVSKAKSINTTWYKLSNYTGQFSGALSPLNDYEFFIRDTKNLEMETTLANSVNL 136
 QY 129 TPF-GSQMVNSGIFRYKALARKIVPF-IRASGSDRVITASAEKFTGFSQAKLADPG 186
 DB 137 NPTGEMNAKRHARFLAQYGYWVENQTSFAVFTSNRCHDTAQYFDG-----LGDK- 190
 QY 187 ANPHQASPVINVII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241
 DB 191 -----FNISLQTISEAESAGANTLSAHSFCPAWDD-DVNDLILKKYDKTGLSGIAK 240
 QY 242 RLEAHLPGVNLDEDVNLMDMCPEDTVARTSDATQLSPFCDLFTHDWIDYDIQSLGK 301
 DB 241 RUNKENKGLNLTSSANTFFAWCAYEINAR-----GYSDICNIETFKDELVRFSYGDQLET 295
 QY 302 YGYGAGNPLGPAQGVGFNELLIALRTSPVQDHTSTNHTLDSNPATPLNATLYADPSH 361
 DB 296 YQQTGPGYDVRSVCANLFNASVKLLKSEVQDQ-----KWLSTFH 337
 QY 362 DNTWYSIFALLGNGTKPLSTTSVESIEETDGYSGASWTVPFAARAYVEMMOCEAEKPL 421
 DB 338 DTDILNLTFTIGIIDQNNTAEHPFMENT--FHRSNVVPQGARVYTEKFCQ--SNDTY 393
 QY 422 VRVLVNDRVVPLHGCVKDLGCKRDDF-----VEGLSFAR 457
 DB 394 VRVINDAVVPIETCTGPGFSCINDFGYAEKRVAGTDFLK 436

RESULT 7

PPA3_YEAST

ID PPA3_YEAST

AC P24031;

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO3 OR YBR092C OR YBR0813.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85037940; PubMed=6093051;
 RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RT "Structural analysis of the two tandemly repeated acid phosphatase
 genes in yeast."
 RL Nucleic Acids Res. 12:7721-7739(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=S288C;
 MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
 RT Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

Query Match 15.8%; Score 391; DB 1; Length 467;

Best Local Similarity 27.0%; Pred. No. 9.4e-24;

Matches 117; Conservative 61; Mismatches 192; Indels 64; Gaps 13;

QY 49 GTYSPFSLADESAISPDPVKGRVTFVQVLSRHGARYPTSSASKAYSAIEAIOKNATA 108

DB 42 GGAGPYFFPGDYGISRLPESCCKQVQVGRHGARYPTSSASKAYSAIEAIOKNATA 101

QY 109 FKGYAFK-TYNYTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 102 FNSLSFLNDYEFFIRDDDDLEMETTFANSNDVNLNPTGEMDAKRAHAREFLAQYGYMFE 161
 QY 153 KIVPFF-IRASGSDRVIAAEKIEGFSQAKLADPGANPHQASPVINVLPEGAGYNNTLD 211
 Db 162 NQTSFFIFAASSERVHDTAQYFIDGL-----GDQFNISLQTSVAMSAGA---NTLS 210
 QY 212 HG-LCTAFSESLGDDVEANFTAVFAPPIRAREALHLPVNLTDVNLMDMCPFDIVA 270
 Db 211 AGNACPGWDE-DANDDILDKYDTTLDIAKRLNKENGLNLSKDANTLFAWCAYELNA 269
 QY 271 RTSDATQLSFCDLFTHDEWIOYDVLQSLGKYGYGAGNPLCPAOCVGFVNLIELARLTHS 330
 Db 270 R-----GYSDVCDIFTEDELVRYSYGODLVSPYQDGPYDMIRSVGANFLNATLKLKQS 324
 QY 331 PVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFFAALGLYNGTKPLSTTSVESIE 390
 Db 325 ETQD-----LKVWLSFTHDTHDILNLTAGIIDDKNLTAETVFPWG 366
 QY 391 ETDGYSASTVPPFAARAYEMMOCAEKEPLVRVLVNDVRVPLHCGGVDKLGRCKRDDF- 449
 Db 367 NT--FKHSWYVPOGARVYTEKQC--SNDTYVRYVINDAVVPVETCSRGPFGFSCINDFY 422
 QY 450 -----VEGLSEAR 457
 Db 423 DYAEKRVAGTDFLK 436

RESULT 8

ID PPAD_YEAST STANDARD; PRT; 468 AA.
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROBABLE ACID PHOSPHATASE YDL024C PRECURSOR (EC 3.1.3.2).
 GN YDL024C OR D2815
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Vissers S.; Urrestarazu L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z48432; CAA88335.1; -
 CC EMBL; Z74072; CAA98583.1; -
 CC SGD; S0002182; YDL024C.
 CC InterPro; IPR000560; -
 CC Pfam; PF00328; acid_phosphat; 1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; Multigene family.
 CC SIGNAL 1 20 BY SIMILARITY
 CC CHAIN 21 468 PROBABLE ACID PHOSPHATASE YDL024C.
 CC ACT_SITE 76 76 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 CC ACT_SITE 338 338 PROTON DONOR (BY SIMILARITY).
 CC CARBOHYD 98 98 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 163 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 193 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 202 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 238 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 251 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 316 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 357 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 391 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 457 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 462 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDF162C CRC64;
 Query Match 15.0%; Score 369; DB 1; Length 468;
 Best Local Similarity 25.3%; Pred. No. 5.4e-22;
 Matches 117; Conservative 60; Mismatches 175; Indels 110; Gaps 14;
 QY 49 GTYSPFSLADESAISDPVPGKGRVTVQVLSRHGARYPTSSKAYSALIEAIQKNATA 108
 Db 43 GGSAPYEFPPANYGIPDTIDPEGRLTQVOMIGRHGYPTRSEAKDIFEVWYKISNTGK 102
 QY 109 FKGYAFKLT-YNYTL-----GADDLTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 103 YEGSLFLNNGYEFFIPDESLEMETTLQNSIDVLPYTGEMNAKRAHAREFLAKYKGLME 162
 QY 153 KIVPFF-IRASGSDRVIAAEKIEGFSQAKLADPGANPHQASPVINVLPEGAGYNNTLD 211
 Db 163 NCTNPFITTSKRIVDTAQYFAEAL-----GDQFNISLQ 197
 QY 212 HGLCTAFSESELG-----DDVEANFTAVFAPPIRAREALHLPVNLTD 253
 Db 198 ----TLSENSSGANTLAAKSSCPNWNNSNANDILMSYSDYLENISDLNDENKGLNLS 253
 QY 254 DEDVNLMDMCPDVTARTSDATQLSFCDLFTHDEWIOYDVLQSLGKYGYGAGNPLGP 313
 Db 254 RKDAALFSWCAEEL-----NAKGSNICDIFSAELIHSYETDITSTFYQNGPGYKLIK 308
 QY 314 AQGVGVN---ELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIF 370
 Db 309 SIGANFNATVKLIROSAH-----LQKWLSTFTHDTHDILNLT 347
 QY 371 ALGLYNGTPLSTTSVESIEETDGYASWTVPFAARAYEMMOCAEKEPLVRVLVNDVRV 430
 Db 348 TAGLIDDTNLTTHVFPFRDHS--YHRSWYVPOGARVYTEKQC--SNDSYVRYVYVNDVAV 403
 QY 431 VPLHCGGVDKLGRCKRDDFVE-----GLSFARSGGNWEEC 465
 Db 404 VPESCSSGPGFSGEGTGYEYAKDLRGVSF-----YEDC 439

RESULT 9

PHYB_ASPNG
 ID PHYB_ASPNG STANDARD; PRT; 479 AA.
 AC P34754;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE B).
 GN PHYB.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.
 RX MEDLINE=93371452; Pubmed=7916610;
 RA Ehrlich K.C.; Montalbano B.G.; Mullaney E.J.; Dischinger H.C. Jr.,
 RA Ullah A.H.J.;
 RT "Identification and cloning of a second phytase gene (phyB) from
 RT Aspergillus niger (ficcum).";
 RL Biochem. Biophys. Res. Commun. 195:53-57(1993).

QY 108 AFKGYAFLKTYNYTL-----GADDLT-PF-GEQOMVNSGIKRYRYKAL--ARKIYPF 157
 Db 109 EYKGLAFLNDWYYPNENYNAETSPYAGLLDAYNHGNDYKARYGHLWNGETVVPF 168
 QY 158 IRASGSDRVIAAEKIEGFOSAKLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTA 217
 Db 169 F-SSGYGRVIETARKFGEF-----FGVNSTNAALNII 201
 QY 218 FESEELG-----DDVEANETAVFAPPRI--ARLEAHLPGVNLDEDDVNLMD 262
 Db 202 SESEVMGADSLPTCDTNDQTTCONLTVOLFOFKAARLNSQPNMNLTAASDVYNLNV 261
 QY 263 MCFDFVARTSDATQSPFCDLTHDEWIQDYDLSLQKYYGYGAGNPLGPAQGVGFVNE 322
 Db 262 MASFELNAR-----PFSMNINAFQDEWVSFGYVEDLVNYYICAGFGDKNMAVAGVYANA 316
 QY 323 LIARLTHSPVQDHTSTNHTLSDNPATFPLNATLYADFSDHNTMWSIFFALGLY--NGTKP 380
 Db 317 SUTLNOGPK-----AGSLFFNFADTNTITPILAAALGVLIIPNEDLP 358
 QY 381 LSTTSVESIEETDGYSAWTVFAARAYVEMMOCEA-----EKEPLRVLVNDRVVPVPLHCC 436
 Db 359 L-----DRVAFGNPISIGNIVPMGHLTIERSCONTALSDGTYRVRLVNEAVLPFND 413

RESULT 11

PPAL_PICPA STANDARD; PRT; 468 AA.
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE ACID PHOSPHATASE PHO1 PRECURSOR (EC 3.1.3.2).
 GN PHO1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E.; Gannon P.M.; Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RL characterization of the gene and its product."
 RL Gene 163:19-26(1995).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE
 CC -!- INDUCTION: BY PHOSPHATE STARVATION.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U28658; AAA85503.1;
 DR HSSP; P34752; 11HP
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 468 ACID PHOSPHATASE PHO1.
 FT ACT_SITE 84 84 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 345 345 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 52690 MW; AE555B8E27718C2C CRC64;

Query Match 14.2%; Score 350; DB 1; Length 468;
 Best Local Similarity 25.2%; Pred. No. 1.8e-20;
 Matches 124; Conservative 79; Mismatches 20; Indels 90; Gaps 21;

QY 6 VLLSIATL--FGSTSGTALGPRGNSHSCDVGGYQCFPEISHLMGTYSPF--FSLADE 60
 Db 10 IILALATLQSVFAVELQHLVGNDRPYQPT--DQYNILRLHGLG-GPYLYGNGWIAAE 67
 QY 61 SAISPDVPKCRVTFVOVLSRHGARYPTSSAKSAYSALIE-AIOKNATAFKGYAFLKTY 119
 Db 68 SEI-----ESCTIDOAHLMRHGERYPSTNVGKLEALYOKLLDADVEVPTGPLSEFFQDY 122
 QY 120 NYTLG---ADDLTPFGEQOMVNSGIKF-----YRYKAL-----ARKIVPFFRAGSDR 165
 Db 123 DYFSDAANWYEQETTKGYSGLNTAFDFTGLRERYDHLINTSEEGKL--SYWAGSOER 180
 QY 166 VIASAEKFIKGFOSAKLADPGANPHOASPVINVIPE--GAGYNNNTLDHGLCT-----A 217
 Db 181 VVDTAKYFAOGFMKSNYTD-----NVEVVALEEKSGQLNSLTARISCPNYSNI 230
 QY 218 FESEELGDDV---EANTAVFAPIRAREHLPGVNLDEDDVNLMDMCPEDTVARTSD 274
 Db 231 YKDGDFPDNIAERED-----RLNTLSPGNITADDIPTIALYCGPELNVGRGE- 278
 QY 275 ATQLSPFCDLTHDEWIQDYDLSLQKYYGYGAGNPLGPAQGVGFVNELIARLTHSPVQD 334
 Db 279 -----SFCDLVLSREALITYALRLDGLGVNNGNPLGKTIGYVYAN----- 321
 QY 335 HTSTNHTLSDNPATFPLNATLYADFSDHNTMWSIFFALGLYNTK-PLSTTSVESIEETD 393
 Db 322 --ATROLLENTAED-PRDYPLYSFSDHDTLLQVFTSLGLFNVTDLPL-----DQLQFOT 373
 QY 394 GYSASWTVPFAARAYVEMMOCC--EAEKEPLRVLVNDRVVPVPLHGGCGVKLGRCKRDDFVE 451
 Db 374 SFKSTEIVPMGARLLTERLLCTVEGEEKYVVRTILNDVAFPLSDCSPGFCPLNDYVS 433
 QY 452 GLSFARSGGNWEE 464
 Db 434 RLEALNEDSDFAE 446

RESULT 12

PPA5_KLULA STANDARD; PRT; 469 AA.
 AC P52289;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE REPRESENTABLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO5.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RA Ferminan E.;
 RL Thesis (1995), University of Salamanca, Spain.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```

Db 97 SVPIDSVSGNPLSFVPTTPVIEANADALSSGRVLFDMGRQFYERYHELFWASTYN 156
QY 158 IRASGSDRVIAAEKFIQFQSAKLADPCANPHQASPVINVIIPGACGYNTLHDGLCTA 217
Db 157 IYTAQORVVDGALWYGYGMF-----GEDVHNFTNVLSENATAGNSLSYNAQPA 209
QY 218 FFESELGDDVEANFTAVFAPPARLEAHLPGVNIITDDEVNLMDCPFDVARTSDATQ 277
Db 210 SDADFTTTPALEAWRVNVMPIRQRLNPFYSNYLNTDNDILNYGICSEYALQ-----D 264
QY 278 LSPFCDLFDHWDIYDYLQSLGKYGYGAGNPLGPAOGVGVNLIARLTHSPVQDHTS 337
Db 265 YSEFCLEFNSVDLNFYEGDLSFSYGMGNSYKMGISFGGAYANSLANS--RSVENNTQ 322
QY 338 TNHTLDSNPATPPLNATLYADSHDNTWVSIFPFGALGLYNGTKPLSTTSVESIEETDGYSA 397
Db 323 -----QVFFAFTHDANIIPVETALGFFDTNTPENPLTSYQVHSHSKA 366
QY 398 SWTVPPAARAYVEMMQCEAEKEPLRVVLNDRVPLHGCG 437
Db 367 SEFVFPAGNLITELFOCEDSKY-VYRHLVNEEVPLSDCG 405

RESULT 14
PPA2_SCHPO
ID PPA2_SCHPO STANDARD; PRT; 463 AA.
AC Q01682;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE THIAMINE-REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
GN PHO4 OR SPBC428.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
RX MEDLINE=91064763; PubMed=2249257;
RA Yang J., Schweingruber M.E.; and Baker S., Mungall K.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CELL WALL.
CC -!- INDUCTION: REPRESSED BY THIAMINE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56939; CAA040258.1; -
CC EMBL; AL034382; CAA22278.1; -
CC PIR; S14119; S14119.
CC InterPro; IPR000560; -
CC Pfam; PF00328; acid_phosphat; 1.
CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC Hydrolase; Glycoprotein; Cell wall; Signal.

```

```

FT SIGNAL 1 18
FT CHAIN 19 463
FT ACT_SITE 69 69
FT ACT_SITE 340 340
FT CARBOHYD 98 98
FT CARBOHYD 104 104
FT CARBOHYD 186 186
FT CARBOHYD 221 221
FT CARBOHYD 251 251
FT CARBOHYD 328 328
FT CARBOHYD 433 433
FT CARBOHYD 439 439
FT CARBOHYD 458 458
FT SEQUENCE 463 AA; 52118 MW; F48EAF8BB6B234A CRC64;

Query Match 12.3%; Score 303.5; DB 1; Length 463;
Best Local Similarity 26.7%; Pred. No. 9e-17;
Matches 113; Conservative 63; Mismatches 184; Indels 63; Gaps 16;

QY 46 HLWGTVS-----PFFSLADESAISPDPKGRVTFVOVLSHRGARYPT--SSASKAYS-- 97
Db 35 HL-GTISVYHEPYF-----NGPTTSPESCAIKQVHLLQRHGSRNPTGDDTATDVSSAQY 88
QY 98 -----LIEAIQKNATAFKGYAFLKTYNTL---GADDLTPFGEQWVNSGIKFYRR 146
Db 89 IDIFQNKLLNGSIPVNFSPENPLYFVKHWTPVKAENADQLSSSGRIELFDLGROVFER 148
QY 147 YKALARKIVPFIASGSDRVIAASAERFIEGFQSAKLADPCANPHQASPVINVIPE--CA 204
Db 149 YLELFDVVDYDINTAAQERVVDSEAWFSYGM-----FGDDMQNKTNF-----IVLPEDDSA 199
QY 205 GYNTLHDGLCTAFEFSELGDD--VEANFTA---VFAPPTRARLEAHL-PGVNLTDEDEVN 259
Db 200 GANSLAMYSCPYVEDNNIDNTEAAHTSWRNVLKPIANRLNKYFDGYNLTVSDVRS 259
QY 260 LMDMCPFDVTARTSDATQSPFCDLFDHWDIYDYLQSLGKYGYGAGNPLGPAQGVGF 319
Db 260 LYICVYETALRDN-----SDFCSLTPSEFLNFEDSDLDYAYWGGPASEWASTLGGAY 314
QY 320 VNELIARLTHSPVDHTSTNHTLDSNPATPPLNATLYADESHDNTWVSIFPFGALGLYNGTK 379
Db 315 VNNLANNL-----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDIT 358
QY 380 PLSTTSVESTEETDGYASASWTVPFAARAYVEMMQCEAEKEPLRVVLNDRVPLHGCGVD 439
Db 359 PEHPLTDKNIFTYSLKTSFVFPFAGNLITELFLC--SDNKYVYRHLVNOQVYPLTDCGYG 417
QY 440 KLG 442
Db 418 PSG 420

RESULT 15
PPAX_CAEEL
ID PPAX_CAEEL STANDARD; PRT; 755 AA.
AC Q09549; Q17843;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE ACID PHOSPHATASE P26C11.1 (EC 3.1.3.2).
GN P26C11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditia; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Matthews P., Lloyd C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.

```

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z47072; CAA87370.1; -
DR EMBL; Z54342; CAA87370.1; JOINED.
DR EMBL; Z54342; CAA91156.1; -
DR EMBL; Z47072; CAA91156.1; JOINED.
DR WormPep; P26C11.1; CE05732.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein: Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
SQ SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.8%; Score 142; DB 1; Length 755;
Best Local Similarity 23.1%; Pred. No. 0.0014;
Matches 94; Conservative 53; Mismatches 132; Indels 128; Gaps 22;
QY 57 LADESAISPDVPKGVKVFQVLSRHGRYPTSSASKAYSALIEAIOKNATAFKGYAFL 116
DB 332 LAIEKSFMEV-DNMKLEFVQTIWHRGR-----SALEGLFPISEKNWTFGGG----- 378
QY 117 KTYNTLTGADLTPEGOQMVNSGKIFYRRY-----KALARKIYVFTIRAGSDRVI 167
DB 379 -----GLGELTPMGSMNNLGTIFRRYVEDQOFLSHRYAAKEI--YIRSTNLNRTI 429
QY 168 ASAEKFIQFQSAKLADPCANPHQASPINVIIPGAGYNNTLD-----HGL- 214
DB 430 ISAMSLLYG-----MFPFGA-----WNI---QGVDYPNDVDWQOQGTTFIPVHVDGID 473
QY 215 -----CTAFEE-----SELGDDVEANFTAVFAPPPIRARLEAHLPGVNLTDDEVVNL 260
DB 474 QCAVAQLCMCRFEQLQEKWAEI-DEVKNATVAMIA--LNRVAAF-----YVWTDGP----- 523
QY 261 MDMCPEDTVARTSDA--TQLSPFCD-LFTHDEWIQYDYLQSLGKYG-----YGAGNP 310
DB 524 -----EKENRYTDAMKQCRNWFNDTMYQQLPWYNEDLYNEAORTYAPFKRFTGNGFNP 577
QY 311 -----LGPAGQGVFNELIARLTHSPVDHTSTNHTLD--SNPATFPLNATLY 356
DB 578 KPSIVDGDIDPOEVSTLOGGPLLNEIFEFEREKIRCVADAENCSDIYKPLKF-----Y 631
QY 357 ADFSHDNTMVSIFFALGLYNGTKPLSTSVESIEETDGYASASWTPVF 403
DB 632 AYSSHDQVYVALLVTLGI-----TDVVKTVGWDGPDTSLSLTIEY 670

Search completed: October 26, 2001, 15:18:05
Job time: 451 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:26 ; Search time 75.85 Seconds
(without alignments)
814.587 Million cell updates/sec

Title: US-09-488-265-31
Perfect score: 2468
Sequence: 1 MGFFVLLSLATLFGSTSGT.....DFVEGLSFARSGGNWEECPA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	76.5	465	3 000092	000092 aspergillus
2	1847	74.8	467	3 090027	Q90027 aspergillus
3	1841	74.6	467	3 093838	Q93838 aspergillus
4	1828	74.1	467	3 09HEQ0	Q9HEQ0 aspergillus
5	1822	73.8	466	3 000100	Q00100 aspergillus
6	1752	71.0	466	3 000085	Q00085 aspergillus
7	1710	69.3	466	3 000096	Q00096 talaromyces
8	1334.5	54.1	487	3 000107	Q00107 thielavia h
9	338.5	13.7	469	3 09Y846	Q9Y846 kluyveromyc
10	334.5	13.6	442	3 074677	Q74677 pichia anqu
11	311.5	12.6	463	3 060172	Q60172 schizosacch
12	204.5	8.3	460	5 09VV72	Q9VV72 drosophila
13	196	7.9	451	11 035217	Q35217 rattus norv
14	196	7.9	467	5 096421	Q96421 drosophila
15	188	7.6	481	11 0922L6	Q922L6 mus musculu
16	188	7.6	487	4 095172	Q95172 homo sapien
17	187	7.6	487	4 09UNW1	Q9UNW1 homo sapien
18	185	7.5	487	4 09UGA3	Q9UGA3 homo sapien
19	181.5	7.4	453	5 096420	Q96420 drosophila

20	180.5	7.3	453	5 09W438	Q9W438 drosophila
21	177.5	7.2	449	13 092170	Q92170 gallus gall
22	168.5	6.8	198	3 09UTX1	Q9UTX1 schizosacch
23	149.5	6.1	468	10 004509	Q04509 arabidopsis
24	144	5.8	274	11 09JJD5	Q9JJD5 mus musculu
25	130	5.3	449	5 019076	Q19076 caenorhabdi
26	128.5	5.2	374	11 09JMG5	Q9JMG5 mus musculu
27	127.5	5.2	381	11 09QXG5	Q9QXG5 mus musculu
28	124.5	5.0	513	2 046334	Q46334 comamonas t
29	118	4.8	380	5 022525	Q22525 caenorhabdi
30	113.5	4.6	381	11 09QXH7	Q9QXH7 mus musculu
31	112.5	4.6	513	2 024719	Q24719 comamonas t
32	110	4.5	683	5 000838	Q00838 leishmania
33	110	4.5	707	5 000839	Q00839 leishmania
34	109.5	4.4	730	5 020826	Q20826 caenorhabdi
35	108.5	4.4	758	4 09P2C1	Q9P2C1 homo sapien
36	108.5	4.4	827	2 047871	Q47871 eubacterium
37	107	4.3	354	5 019709	Q19709 caenorhabdi
38	106	4.3	763	2 09R6Y7	Q9R6Y7 anabaena sp
39	103.5	4.2	1013	14 09ILX9	Q9ILX9 retroperito
40	103	4.2	421	4 09UIG6	Q9UIG6 homo sapien
41	103	4.2	428	4 09NPH0	Q9NPH0 homo sapien
42	103	4.2	452	5 019175	Q19175 caenorhabdi
43	102.5	4.2	1225	5 020330	Q20330 caenorhabdi
44	101.5	4.1	642	1 028446	Q28446 methanobact
45	101	4.1	408	5 019461	Q19461 caenorhabdi

ALIGNMENTS

RESULT 1

000092 ID 000092 PRELIMINARY: PRT: 465 AA.

AC 000092;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).

GN PHYA.

OS Aspergillus fumigatus (Sartorya fumigata),
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5085;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.

RC STRAIN-ATCC 34625;

RX MEDLINE-97288063; PubMed-9143104;

RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
RT phytase from the fungus Aspergillus fumigatus";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).

CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.

CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMATIC ACTIVITY.

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

DR EMBL; U59804; AAB96872.1; -;
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid.phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.

FT CHAIN 27 465 3-PHYTASE A.
 FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
 FT SIMILARITY).
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 359 359 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 30 39 BY SIMILARITY.
 FT DISULFID 70 412 BY SIMILARITY.
 FT DISULFID 213 463 BY SIMILARITY.
 FT DISULFID 262 280 BY SIMILARITY.
 FT DISULFID 434 442 BY SIMILARITY.
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 465 AA; 50836 MW; 86FC1D9058C9B2C9 CRC64;

Query Match 76.5%; Score 1888; DB 3; Length 465;
 Best Local Similarity 76.9%; Pred. No. 4.5e-144;
 Matches 359; Conservative 35; Mismatches 71; Indels 2; Gaps 2;

QY 1 MGVEVLLSIATLFGSTGTALPGRNHSCDVTGQYQCFPEISHLWGTSPFFSLADE 60
 DB 1 MVTLFLLSAAYLLSCRVSAAEPSSAG-SKSCDVTDLGQYQCFPEISHLWGTSPFFSLADE 59
 QY 61 SAISPDPVKRCRTVFQVLSRHGARYPTSSASKAYSALEIAOKNATAPKGYAFELKTYN 120
 DB 60 LSVSSKLKPKDCRTTLVQVLSRHGARYPTSSASKAYSALEIAOKNATAPKGYAFELKTYN 119
 QY 121 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 120 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 179
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPTR 240
 DB 180 KLADPGAT-NRAAPALSVIIPESETFNNTLDHGVTCKPEASQGLGDEVAANFTALFAPDIR 238
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSFPFCDLTHDEWIDYDLSQLG 300
 DB 239 ARAEKPLPGVTLTDEDEVVNLMDMCPFDVTARTSDATQLSFPFCDLTHDEWIDYDLSQLG 298
 QY 301 KYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 299 KYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 358
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEKEP 420
 DB 359 HDNSMVSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCKSEKEP 418
 QY 421 LVRVLVNDRVVPLHCGVVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467
 DB 419 LVRVLVNDRVVPLHCGVVDKLGCRKDDFVEGLSFARSGGNWEECEFS 465

RESULT 2
 Q90U27

ID Q90U27 PRELIMINARY; PRT; 467 AA.
 AC Q90U27;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (EC 3.1.3.8).
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Hongning W., Qi W., Jing X.;
 RT "PCR, cloning and characterization of the phytase (phyA) gene of

RT Aspergillus niger (China Strain).";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF218813; AAF25481.1; --
 DR HSSP; P34752; 1IHP.
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Signal; Lyase; Hydrolase.
 FT SIGNAL 1 19 POTENTIAL.
 SQ SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 74.8%; Score 1847; DB 3; Length 467;

Best Local Similarity 74.1%; Pred. No. 9.2e-141;

Matches 346; Conservative 45; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALPGRNHSCDVTGQYQCFPEISHLWGTSPFFSLADE 60
 DB 1 MGVSALLPLYLGLSGVTSLGAVPASNRQSTCDVTGQYQCFSETSLWGOYAPFFSLANE 60
 QY 61 SAISPDPVKRCRTVFQVLSRHGARYPTSSASKAYSALEIAOKNATAPKGYAFELKTYN 120
 DB 61 SAISPDPVACGRTVFAQLSRHGARYPTDSKGYKYSALIEEQONATTFDGYAFELKTYN 120
 QY 121 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 121 YSLGADDLTPFGEQELVNSGKIFRYQYSLRNIIPFIRSSGSRVIAASKKPIEGFQST 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPTR 240
 DB 181 KLKDPRAQPCQSSPKIDVISEASSNNLTDPGTCVTFEDELADAVEANFTATFVPTIR 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSFPFCDLTHDEWIDYDLSQLG 300
 DB 241 QRLNDLSGSLTDEVTYLYMDMCSFDTISTVTDKLSFPFCDLTHDEWIDYDLSQLK 300
 QY 301 KYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISLLFALGLYNGTKPLSTTIVQNIITQDGFSSAWTVPFASRLYVEMMOCEAEKEP 420
 QY 421 LVRVLVNDRVVPLHCGVVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467
 DB 421 LVRVLVNDRVVPLHCGVVDKLGCRKDDFVEGLSFARSGGNWEECEFA 467

RESULT 3

Q93838
 ID Q93838 PRELIMINARY; PRT; 467 AA.
 AC Q93838;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 GN PHYTASE.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA STRAIN-SK-57;
 RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
 RT "Phytase having high-affinity for phytic acid."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB022700; CAB19824.1; --
 DR HSSP; P34752; 1IHP.
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid_phosphat; 1.

DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543EDC265C CRC64;

Query Match 74.1%; Score 1841; DB 3; Length 467;
 Best Local Similarity 73.2%; Pred. No. 2.8e-140;
 Matches 342; Conservative 49; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRGNHSCDVTVDGGYQCFFPEISHLWGYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSTCDTVDGGYQCFFSETSHLWGYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATFAFGKYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFQVLSRHGARYPTDSKGYKYSALIEEIQONATTEGKYAFKLTYN 120
 QY 121 YTLGADLLTPFGQOMVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 YSLGADLLTPFGQELVNSGVKFYQRYESLNRNIVPFFIRSSGSRVIAASGNKFIQFQST 180
 QY 181 KLADPCANPHQASPVNVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGQSSPKIDVIVSEASTSNNTLDPGCTCTVFEDSELADDEANFTATVPPIR 240
 QY 241 ARLEAHLPGVNLTDVDVNLMDMCPDFTVARTSDATOLSPFCDLTFHDEWYQYDLSLG 300
 DB 241 ORLENDLSGVSLTDEVTYLMDCSFDITSTVDTKLSPPCDLTFHEWINYDYLQSLN 300
 QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYGHCAGNPLGPTQGVGYANELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEERTDCYSASWTVPFAARAYVEMMOCEAKEP 420
 DB 361 HNGIISILFALGLYNGTKPLSTTAENITQDGFSSARTVPFAARMYVEMMQCQSEEP 420
 QY 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEFVGLSFARSGGNWEECF 467
 DB 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVKGLSFARSGGDWAECEFA 467

RESULT 4

ID Q9HEQ0 PRELIMINARY; PRT; 467 AA.
 AC Q9HEQ0;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PHYTASE.
 OS Aspergillus ficuum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L., An L., Wang Y., Yuan X.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY013315; AAG40885.1; -;
 SQ SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match 74.1%; Score 1828; DB 3; Length 467;
 Best Local Similarity 73.0%; Pred. No. 3.2e-139;
 Matches 341; Conservative 49; Mismatches 77; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRGNHSCDVTVDGGYQCFFPEISHLWGYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSTCDTVDGGYQCFFSETSHLWGYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATFAFGKYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFQVLSRHGARYPTDSKGYKYSALIEEIQONATTEGKYAFKLTYN 120

QY 121 YTLGADLLTPFGQOMVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 YSLGADLLTPFGQELVNSGVKFYQRYESLNRNIVPFFIRSSGSRVIAASGNKFIQFQST 180
 QY 181 KLADPCANPHQASPVNVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGQSSPKIDVIVSEASTSNNTLDPGCTCTVFEDSELADDEANFTATVPPIR 240
 QY 241 ARLEAHLPGVNLTDVDVNLMDMCPDFTVARTSDATOLSPFCDLTFHDEWYQYDLSLG 300
 DB 241 ORLENDLSGVSLTDEVTYLMDCSFDITSTVDTKLSPPCDLTFHEWINYDYLQSLN 300
 QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYGHCAGNPLGPTQGVGYANELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HNTWYSIFFALGLYNGTKPLSTTSVESIEERTDCYSASWTVPFAARAYVEMMOCEAKEP 420
 DB 361 HNGIISILFALGLYNGTKPLSTTAENITQDGFSSARTVPFAARMYVEMMQCQSEEP 420
 QY 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEFVGLSFARSGGNWEECF 467
 DB 421 LVRVLNDRVPLHGGVVDALGRCTRDSFVKGLSFARSGGDWAECEFA 467

RESULT 5

ID 000100 PRELIMINARY; PRT; 466 AA.
 AC 000100;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE PRECURSOR (EC 3.1.3.6) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE PHOSPHOHYDROLASE).
 DE Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 116.46;
 RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B., Broger C., van Loon A.P.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHITATE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U60412; AAB58465.1; -;
 DR HSSP; P34752; I1HP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 466 3-PHYTASE.
 FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 31 40 BY SIMILARITY.
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).

Sat Oct 27 15:25:56 2001

us-09-488-265-31.rspt

FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51055 MW; F2AECECIAF7C22C4 CRC64;

Query Match 73.8%; Score 1822; DB 3; Length 466;
 Best Local Similarity 72.3%; Pred. No. 9.6e-139;
 Matches 337; Conservative 46; Mismatches 88; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNSHSDCTVDGQYQCFPELHKWGLYAPYFSLQDE 60
 DB 1 MGVEVLLSIATLFGSTGTALGPRGNSHSDCTVDGQYQCFPELHKWGLYAPYFSLQDE 60

QY 61 SAISDPVPGKRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFKGYAFLKTYN 120
 DB 61 SPFLDVPDCHITFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQQMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 YSMGSENLNPFGRNQLDGLGAQFYRYDTRHINPFVRAADSRVHSAEKFEVGFQNA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNVLDHGLCTAPESELGDDVEANFTAVFAPPIR 240
 DB 181 ROGDHHPANPHQASPVINVIPEGAGYNNVLDHGLCTAPESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVVDVNLMDMCPDVTARTSDATQLSPFCDLTHDQWYDYLQSLG 300
 DB 241 KRLEADLPVQLSADDDVNLMDMCPDVTARTSDATQLSPFCDLTHDQWYDYLQSLG 300

QY 301 KYGYGAGNPLGPAQGVGVFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
 DB 301 KYGYGAGNPLGPAQGVGVFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360

QY 361 HDNTMVSIFALGLYNGTKPLSTTSVESTETDGYASMTVPFAARAYVEMMOCEAEK 420
 DB 361 HDNTMVSIFALGLYNGTKPLSTTSVESTETDGYASMTVPFAARAYVEMMOCEAEK 420

QY 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDFVGLSFAAGNNAECF 466
 DB 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDFVGLSFAAGNNAECF 466

RESULT 6
 ID 000085 PRELIMINARY; PRT; 466 AA.
 AC 000085;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHA.
 GN PHA.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A1;
 RC MEDLINE=97177792; PubMed=9025298;
 RX Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
 RA "The phytase subfamily of histidine acid phosphatases: isolation of
 RT genes for two novel phytases from the fungi Aspergillus terreus and
 RT Myceliophthora thermophila";
 RL Microbiology 143:245-252(1997).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 CC 2.5 TO 7.5, WITH MAXIMAL ACTIVITY AT PH 5.5. ALSO ACCEPT 4-
 CC NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO
 CC MORE ACIDIC PH VALUES.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59805; AAB52507.1; -;
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; -;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 466 3-PHYTASE A.
 FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY
 FT ACT_SITE 83 83 SIMILARITY).
 FT ACT_SITE 361 361 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT DISULFID 31 40 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51093 MW; 21CDB5959C96A866 CRC64;

Query Match 71.0%; Score 1752; DB 3; Length 466;
 Best Local Similarity 69.3%; Pred. No. 4.2e-133;
 Matches 323; Conservative 56; Mismatches 87; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNSHSDCTVDGQYQCFPELHKWGLYAPYFSLQDE 60
 DB 1 MGVEVLLSIATLFGSTGTALGPRGNSHSDCTVDGQYQCFPELHKWGLYAPYFSLQDE 60

QY 61 SAISDPVPGKRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFKGYAFLKTYN 120
 DB 61 SPFLDVPDCHITFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQQMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 YSLDSEELTPFGNQLDGLGAQFYRYDTRHINPFVRAADSRVHSAEKFEVGFQNA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNVLDHGLCTAPESELGDDVEANFTAVFAPPIR 240
 DB 181 RQDDHHPANPHQASPVINVIPEGAGYNNVLDHGLCTAPESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVVDVNLMDMCPDVTARTSDATQLSPFCDLTHDQWYDYLQSLG 300
 DB 241 QREADLPVQLSADDDVNLMDMCPDVTARTSDATQLSPFCDLTHDQWYDYLQSLG 300

QY 301 KYGYGAGNPLGPAQGVGVFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
 DB 301 KYGYGAGNPLGPAQGVGVFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360

QY 361 HDNTMVSIFALGLYNGTKPLSTTSVESTETDGYASMTVPFAARAYVEMMOCEAEK 420
 DB 361 HDNTMVSIFALGLYNGTKPLSTTSVESTETDGYASMTVPFAARAYVEMMOCEAEK 420

QY 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDFVGLSFAAGNNAECF 466
 DB 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDFVGLSFAAGNNAECF 466

RESULT 7
 ID 000096 PRELIMINARY; PRT; 466 AA.
 AC 000096;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHA.
 GN PHA.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A1;
 RC MEDLINE=97177792; PubMed=9025298;
 RX Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
 RA "The phytase subfamily of histidine acid phosphatases: isolation of
 RT genes for two novel phytases from the fungi Aspergillus terreus and
 RT Myceliophthora thermophila";
 RL Microbiology 143:245-252(1997).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

QY	239	IFARLEAHLPGVNLITDEDVNLMDCPFDTVARTSDATQLSPFCDLFTHDENIQYDYLQS	298
DB	236	ILEIKTKHLPGVDLAVSDVPLMLDCLPFETIARNHTDT-LSPFCALSTQEOWAYDYQS	294
QY	299	LKYYGYGAGNPLCPAGCVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYAD	358
DB	295	LKYYGNGGNGPLCPAGCVGFVNELIARMTHTSPVQDITVTNHTLDSNPATPLNATLYAD	354
QY	359	FSHDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK	418
DB	355	FSHDNTMTSIFAALGLYNGTAKLSTTEIKSIEETDGYSAWTVPGFRAYIEMMQCDDSD	414
QY	419	EPLRVVLNDRVYPLHCGGYVDKLGRCRRDDFVEGLSFARSGGNWECFCA	467
DB	415	EPVVRVLNDRVYPLHCGEVDLSGRCKRDDFVRGLSFARQGGNWEGCYA	463
RESULT	8		
ID	000107	PRELIMINARY;	PRT; 487 AA.
AC	000107;		
DT	01-JUL-1997	(Tremblrel. 04, Created)	
DT	01-JUL-1997	(Tremblrel. 04, Last sequence update)	
DT	01-MAY-2000	(Tremblrel. 13, Last annotation update)	
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3- DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE GN. PHA.		
OS	Thielavia heterothallica.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
OC	Sordariales; Chaetomiaceae; Thielavia.		
OC	NCBI_TaxID=78579;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97177792; PubMed=9025298;		
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,		
RA	van Loon A.P.G.M.;		
RT	"the phytase subfamily of histidine acid phosphatases: isolation of genes for two novel phytases from the fungi Aspergillus terreus and Myceliophthora thermophila.";		
RL	Microbiology 143:245-252(1997).		
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE.		
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO- INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.		
CC	-1- SUBCELLULAR LOCATION: SECRETED.		
CC	-1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF 3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO MORE ACIDIC PH VALUES.		
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.		
DR	EMBL; U59806; AAB52508.1; ..		
DR	HSSP; P34752; 1IHP.		
DR	InterPro; IPR000560; ..		
DR	Pfam; PF00328; acid.phosphat. 1.		
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.		
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.		
KW	Hydrolase; Glycoprotein; Signal.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	? 487	3-PHYTASE A.
FT	DOMAIN	267 270	POLY-SER.
FT	DOMAIN	423 433	POLY-GLY.
FT	ACT_SITE	75 75	REQUIRED FOR BINDING SUBSTRATE (BY SIMILARITY).
FT	ACT_SITE	76 76	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE	368 368	PROTON DONOR (BY SIMILARITY).
FT	DISULFID	26 35	BY SIMILARITY.
FT	DISULFID	64 421	BY SIMILARITY.
FT	DISULFID	208 485	BY SIMILARITY.
FT	DISULFID	260 289	BY SIMILARITY.
FT	DISULFID	456 464	BY SIMILARITY.
FT	CARBOHYD	165 165	N-LINKED (GLCNAC...) (POTENTIAL).
FT	CARBOHYD	200 200	N-LINKED (GLCNAC...) (POTENTIAL).

[illegible]

[illegible]

1111

SQ SEQUENCE 442 AA; 49370 MW; 7087D91A85B05C31 CRC64;

Query Match 13.6%; Score 334.5; DB 3; Length 442;
Best Local Similarity 25.8%; Pred. No. 8.3e-19;
Matches 106; Conservative 63; Mismatches 159; Indels 83; Gaps 16;

QY 66 DVPKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIQK-NATAFKGKYAFKLTNYNTLG 124
DB 55 DTPPHCEIEQAQLFMRHGERFPFKSSGQYKKEYDKLKKANITDYKGPLAFIEDLEYFVP 114
QY 125 ADDLTPFGQOMVNSGI-----KF-----YRYKAL--ARKIVPPIRASGSDRVIASAEKFI 174
DB 115 DSDNYELETTRGLYGLLNAFAKGYLRYRYSLDVTSVLPPIFAAS-EDRVVDVARSFG 173
QY 175 EGFOSAKLADPGANPHQAQSPVINVI-----IPEGAGYNNVLDHGLC-----TAFEESELGDD 226
DB 174 REFFGPDYA-----TSCSQVNNETTSKGANALTTKDN--CPTYNSFSYDYSFGDE 223
QY 227 V---EANTFVAPPPIRARLEAHLPGVNLTDDEVDVMDMCPDPTVARTSDATQLSPFCD 293
DB 224 IFQREAD-----RLNELSPGFNITADDIITMGTYCAVETNVKH-----SSFCD 267
QY 284 LETHDEWIOYDLSGLGKYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLD 343
DB 268 ALSREAFIALQVNNVDVTKYQFGPGYMSAVAGGVYAN----- 305
QY 344 SNPATFPL---NATLYADSHDNTWVSIFALGLYNGTKPLSTTSVESTTEEDGYSASWT 400
DB 306 ---ATAKLQEDGKWLWFSHNDLLNYTALGLITDTE---LGTEDYDFHRSFTSEL 358
QY 401 VPFARAYVEMQCEAEKPLVRLVNDVRVPLHGGGVGDKLGRCKRDDPVE 451
DB 359 VFGQARLIEKLNCC--SDTSFVRTILNDKVYVPGCGSSGPGYSCPLEYLD 407

RESULT 11
O60172

ID O60172 PRELIMINARY; PRT; 463 AA.
AC O60172;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE PROBABLE ACID PHOSPHATASE (EC 3.1.3.2).
GN SPBC21H7.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC EMBL; AL023286; CAAL8863.1; -.
DR HSP; P34755; IQFX.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid phosphat; 1.
DR PROSITE; PS00616; HIS-ACID-PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS-ACID-PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein.
FT ACT_SITE 69 69 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 340 340 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 463 AA; 52758 MW; 6C41AF422C6D624A CRC64;

Query Match 12.6%; Score 311.5; DB 3; Length 463;
Best Local Similarity 25.3%; Pred. No. 6.4e-17;
Matches 113; Conservative 69; Mismatches 197; Indels 67; Gaps 15;

QY 46 HLMGTYS-----PFPSIADESAISPDPKGRVTFVQVLSRHGARYPTS--SASKAYS-- 97
DB 35 HL-GTILSPYHEPYFDGLDSA-----FPETCEIQOVHLLQRHGSRNPTGDTVATDVYSSQY 88
QY 98 -----LLEATQKNATAFKGKYAFKLTNYNTL---GADDLTPFGQOMVNSGIKPYRR 146
DB 89 LNNFQEKLLNGSPVNFSPENPLCFIKQWTPVIDAENADQLSSRGLRLEFLDLGRQLYQR 148
QY 147 YKALARKIYVPIRASGSDRVIASAEKFIQFQSAKIADPGANPHQAQSPVINVLIPG--A 204
DB 149 YKLFDSYVVDINTAQERVVESAKWFTYGLFGDKKWE-----KTNFLLISEKAA 199
QY 205 GYNTLHDGLCTAFESEL-----GDDVEANFTAVFAPPPIRARLEAHL--PGVNLTDDEVN 259
DB 200 GANSLSWYNACPVKKNFHNKATDAAHAVWRNFIPIVNRKAYDFDSSYKLTINDVRS 259
QY 260 LMDMCPDPTVARTSDATQLSPFCDLETHDEWIOYDLSGLGKYGYGAGNPLGPAQGVGF 319
DB 260 LFYICYEIAIKDH-----SDFCSIFTPSEFLNFEDSDLDQAYGGPVSEWASTILGGAY 314
QY 320 VNELIARLTHSPVQDHTSTNHTLDSNPATFPPLNATLYADFSDHNTWVSIFALGLYNGTK 379
DB 315 INNLADSL-----RNVTNPFDRK-----VFLAFTHDSNIIPVEAALGFPPDIT 358
QY 380 PLSTTSVESTTEEDGYSASWTVPFAARAYVEMQCEAEKPLVRLVNDVRVPLHGGGVGD 439
DB 359 PQNPLPTDKNIVTYSQKTSSTSFVFNAGNLITLFFC--SDSKYVYVRLHVNQVPLIDCGVG 417
QY 440 KLGR-----CKRDDFVEGLSFARSGGN 461
DB 418 PSGTSGDLCEQLQAYLNSPIRANSTSN 443

RESULT 12

Q9VV72
ID Q9VV72 PRELIMINARY; PRT; 460 AA.
AC Q9VV72;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE MIPPI PROTEIN.
GN MIPPI OR CG4123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov E.M.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong E., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003527; AAF49450.1;
 DR FlyBase: FBgn0026061; Mippl.
 DR InterPro: IPR000560;
 DR Pfam: PF00328; acid_phosphat; 1.
 SQ SEQUENCE 460 AA; 52685 MW; F00B25718E40807D CRC64;

Query Match 8.3%; Score 204.5; DB 5; Length 460;
 Best Local Similarity 23.9%; Pred. No. 2.7e-08;
 Matches 106; Conservative 64; Mismatches 171; Indels 103; Gaps 22;

QY 66 DVPK-----GCRVTFVQVLSRHGARYTSS-----ASKAYSALIEAIQKNATAFKKG-----112
 DB 46 DIDKQVLPGCCQPKQWIFRHGTRLPKSMINKASRV-AELRLDIINNYQVARTKPTD 104
 QY 113 -----YAFKYNVYTLGADDTLPGEQMVNSGKIFRYKALARKIVP-FIRASGSD- 164
 DB 105 ALQOTDLIAIKLWKN-----SSITPDMEEYLTAAQYEDLRTGAKLYRYPTVLTANYNDT 161
 QY 165 ---RVTAEEKFEG-FQSAKLADPCANPHO-----ASPVINV-IIPGAGYNNLT 210
 DB 162 YQRTTESKFAEAGLFGSONAAHVEIPKQDILLRPDYCSFKNVNVKDEGSEYK-- 219
 QY 211 DHGLCTAFEESELGDGVANFTAVFAPPIRARLEAHPG--VNLTDEDVNNLMDMCPEDT 268
 DB 220 -----FHQSKLYNTLAD-----ISTRL-----GFLYTLLEADIKLMDYMCRYE- 258
 QY 269 VARTSATQLSPCDLFTDHWIYQDYLOSGLKYYGYGAGNPLGPAQGVFNELIARLT 328
 DB 259 --QAWNVDNRNSVMCGAFLEQITVFYLEDLYYYGSGYGFPENAHLNRLVODLLTHLS 316
 QY 329 HSPVQDHTSNHTLDSNPATFPLNATFLYADESHDNMTWISFFALGLYNGTKPLSTTSVES 388
 DB 317 -NPVSPH-----VVAHFGHSTGLLLTALGKQDDIKLRADNYDS 356
 QY 389 IETDGYASWTVFAARAYVEMMOCEA--EKEPLRVLVNDRVPLHCGGVKDLGRCKR 446
 DB 357 L-TSRRWKSLIDPFAANFVAVKYDCPADLDREKVV-FELNQAVQLDWCVS---GLCKW 411
 QY 447 DDFVE-----GLSFARSGG 460
 DB 412 SDVLEKYTKTIADACGEYCYRTGG 435

RESULT 13
 ID 035217 PRELIMINARY; PRT; 451 AA.
 AC 035217;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-AUG-1999 (TREMBlrel. 11, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE HEPATIC MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE (EC 3.1.3.-)
 DE (MIPP).
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE OF 11-451 FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE=LIVER;
 RX MEDLINE=98028556; PubMed=9359836;
 RA Craxton A., Caffrey J.J., Burkhardt W., Saifany S.T., Shears S.B.;
 RT "Molecular cloning and expression of a rat hepatic multiple inositol
 RL polyphosphate phosphatase."
 RL Biochem. J. 328:75-81(1997).
 RN [2]
 RP CATALYTIC ACTIVITY.
 RX MEDLINE=91358435; PubMed=1653239;
 RA Nogimori K., Hughes P.J., Glennon M.C., Hodgson M.E., Putney J.W. Jr.,
 RA Shears S.B.;
 RT "Purification of an inositol (1,3,4,5)-tetraakisphosphate 3-phosphatase
 RT activity from rat liver and the evaluation of its substrate
 RT specificity."
 RL J. Biol. Chem. 266:16499-16506(1991).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INOSITOL 1,3,4,5-
 CC TETRAKISPHOSPHATE TO INOSITOL 1,4,5-TRIPHOSPHATE AND INOSITOL
 CC 1,3,4,5,6-PENTAKISPHOSPHATE TO INOSITOL 1,4,5,6-TETRAKISPHOSPHATE
 CC AND THEN TO INOSITOL 1,4,5-TRISPHOSPHATE. DEPHOSPHORYLATES
 CC INOSITOL HEXAKISPHOSPHATE WITHOUT SPECIFICITY TOWARDS A PARTICULAR
 CC PHOSPHATE GROUP.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY AND LIVER.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: AF012714; AAC53453.1; --
 DR InterPro: IPR000560; --
 DR Pfam: PF00328; acid_phosphat; 2.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN-1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; FALSE_NEG.
 KW Hydroxylase; Glycoprotein; Phosphorylation; Multifunctional enzyme.
 FT ACT_SITE 59 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT MOD_RES 65 65 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 146 146 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 176 176 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 201 201 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 218 218 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 391 391 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 412 412 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 435 435 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 447 447 PHOSPHORYLATION (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 3 3 P -> H.
 FT CONFLICT 327 327 S -> P (IN AAC53453).
 FT SITE 448 451 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 451 AA; 51592 MW; EB1C05512A03020B.CRC64;

Query Match 7.9%; Score 196; DB 11; Length 451;
 Best Local Similarity 21.3%; Pred. No. 1.3e-07;
 Matches 91; Conservative 72; Mismatches 187; Indels 78; Gaps 17;

QY 52 SPFTS-----LADESA--ISPDVPGK-CRVTFFVQVLSRHGARYPTTSASKAYS 96
 DB 14 SPYGTGTRYEDVNPWLLGDPVAPRPDELLAGTCTPVQLVALIRHGTTRYPTTKIRKL 73
 QY 97 ALIEAIQKNATAFKGYAFKTYNTLTGADDTLPFGQQMVNSGKIFRYKALARKIVP 156
 DB 74 QLQGLLOTRESVDGSGSRVAAALDQWPLMYDD---WMDGQVLEKGRQMDROLALALFP 130
 QY 157 --FIR-----ASGSDRVIAAEKFIQFQSAKLADPCANPHQASPVINVIPEGAG 205

```

Db 131 DLFCRENYGRRLITSSKHRCDSSAAFLQGLW--QHYHPLPP-----PDVSDMECDPPR 184
QY 206 YNNLT-----DHGLCTAFEESELGDDVEANFTAVF-----APPIRARLE-----AHLPG 249
Db 185 VNDKLMRFEDH--CEKFL-----TEVERNATALLYHVEAEKFTGEMOTVLKKAATLQVPV 237
QY 250 VNLTDEVDVNLMDMCPDFTVARTSDATOLSPFCDLTFHDEWIQYDYLQSLGKYGYGAGN 309
Db 238 NNLNADLIQVAFETCSFDLAIQGVH-----SPWCDVFDVDAKVLVLELNDLKQYWKRSYGY 293
QY 310 PLGPAQGVCFVNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMYSIF 369
Db 294 AINSRSSCNLFQDIFLHLDKAVEQKORSQ-----PVSSVILQFGHAETLPL 342
QY 370 FALGLYNGTKPLSTTSVESIEETDGYSSASWTVPFAARAYVENMOCEAKEP-----LVRVL 425
Db 343 SLMGFYKDEPLTAYNFEEQVIRE--FRSGHIVPYASNLFIYVHCEDAOTPOEKFOIOML 401
QY 426 VNRDVVPL 433
Db 402 LNEKVLPL 409

RESULT 14
ID O96421 PRELIMINARY; PRT; 467 AA.
AC O96421;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 1.
GN MIPPI OR CG4123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
RA Reynolds P.R.
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF046913; AAD02436.1; -.
DR FlyBase; FBgn0026061; Mippl.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
SQ SEQUENCE 467 AA; 53602 MW; 1A0571B9958B2632 CRC64;

```

```

Query Match 7.9%; Score 196; DB 5; Length 467;
Best Local Similarity 22.8%; Pred. No. 1.3e-07;
Matches 103; Conservative 66; Mismatches 172; Indels 110; Gaps 21;

```

```

QY 66 DYPK-----GCRVTFVQVLSRHGARYPTSS-----ASKAYSALIEALOKNATAFKGK---- 112
Db 46 DIDQYLVPGCQOPQOPQIFHRHGTLPLPKSMINKASRV--AELRDLINNYQVARTKPTD 104
QY 113 -----YAFKTYNTLGGADLTTPGEOQVNSGKIFYRRYKALARKIVPFI----- 158
Db 105 ALCOYDLIAIKLWKN-----SSITPDMEVLTAGQVEDLGTAKLQRYVPTLTANYNDT 161
QY 159 -----RASSDRVIASAEKFIQF--FOSAKIADPGANPHQ-----ASPFINV--IIPBG 203
Db 162 YYQFRHTDTQRTTSFKAFAGLFGSQNAHPVEIPKQDILLRPDYDCSSFKNVYKDEG 221
QY 204 AGYNTLHGLCTAFEESELGDDVEANFTAVFAPIRARLEAHLPG--VNLTDEVDVNL 261
Db 222 SEYK-----FHQSKLYNDTLAD-----ISTRN-----GFLYTLLEADIKLMY 259
QY 262 DMCPTDVTARTSDATOLSPFCDLTFHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFN 321
Db 260 DMCRYE---QAWNVRNNSVWCAGAFLEQITVFYLEDLKYGYGYPENAHNLNCRVLQ 316

```

```

QY 322 ELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMYSIFPALGYNKTKPL 381
Db 317 DLTLHUS--NPVSPH-----VVXHFHSTGLLTLLTALGXIKKODIKL 356
QY 382 STTSVESIEETDGYSSASWTVPFAAR--AYVENMOCEAKEPVLVRVLDNRVPLHGCGVD 439
Db 357 RADNYDSL--TSRWKSSLLIDPFAANFVAVKYDLPADLDREKVV--FELNQCAVQDWCSCV- 413
QY 440 KIGRCRKRDFVE-----GLSFARSGG 460
Db 414 --GLCKRSDVLEKYKTIADADCGEYICRTGG 442

RESULT 15
ID Q922L6 PRELIMINARY; PRT; 481 AA.
AC Q922L6;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
GN MINPPI OR MIPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
RA Reynolds P.R.
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF046908; AAD02434.1; -.
DR MGD; MGI:1336159; Minppl.
DR InterPro; IPR000560; -.
DR InterPro; IPR000886; -.
DR Pfam; PF00328; acid_phosphat; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN.1.
SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;

```

```

Query Match 7.6%; Score 188; DB 11; Length 481;
Best Local Similarity 20.8%; Pred. No. 6.2e-07;
Matches 99; Conservative 70; Mismatches 196; Indels 112; Gaps 20;

```

```

QY 11 AILFGSTGTALGPRGNSHSCDVTGCGYOCFPEISHLWNTYSPFFSLAD--ESA-----I 63
Db 21 AALLSFARCSLPGRGD-----PVASVL-----SPYFGTKTRYEDANPWLIV 62
QY 64 SPDVPK-----GCRVTFVQVLSRHGARYPTSSASKAYSALIEALOKNATAFKGYAF 115
Db 63 DPVAPRRDPELLAGTCTPVQVVALIRHGTRYPTTKIRKLQLOGLLQTRSDGGSQVA 122
QY 116 LKTYNYTLGADLTTPGEOQVNSGKIFYRRYKALARKIVP--FIR-----ASGSD 164
Db 123 AALAEWPLWYGD---WMDQGLVEKGRQDMRQLALRALAALPDLFSRENVDRLRLITSSKH 179
QY 165 RVIASAEKFIQFQSAKLADPGANPHQAS-----PVINVIPEGAGYNTLHGLCTAF 218
Db 180 RCVDSSAFLQGLW--QHYHPLPPDPVSDMECGPPRINDKL-----MRFFDH--CEK 229
QY 219 EESELGDDVEANFTAVFAPPIRARLEAHLPG-----VNLTDEVDVNL- 260
Db 230 L-----TDVERNETALY-----HVEAFYTGEMQKVLKKAATLOVPNLSNADIQVA 278
QY 261 DMCPTDVTARTSDATOLSPFCDLTFHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGF 320
Db 279 FTFCSFDLAIKGVH-----SPWCDVFDVDAKVLVLELNDLKQYWKRSYGYT INSRSSCNLF 334
QY 321 NELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMYSIFALGYNKTKP 380
Db 335 QDIFLHLDKAVEQKORSQ-----PVSSPVLQFGHAETLPLLSLMLGKYFKDKPE 383

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:04 ; Search time 57.24 Seconds

(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-31

Perfect score: 2468

Sequence: 1 MGVEVLLSTATLFGSTSGT.....DFVEGLSFARSGGNWEECA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2468	100.0	467	20	AA43170
2	2468	100.0	467	21	AA20527
3	2456	100.0	467	21	AA69569
4	2456	99.5	467	21	AA20534
5	2451	99.3	467	21	AA20533
6	2448	99.2	467	21	AA20524
7	2448	99.2	467	21	AA69566
8	2367	95.9	467	21	AA20532
9	2362	95.7	467	20	AA43169
10	2362	95.7	467	21	AA20526
11	2362	95.7	467	21	AA20531

12	2362	95.7	467	21	AA69568	Mutant phytase-1.
13	2322	94.1	441	21	AA20523	Consensus phytase
14	2314	93.8	467	20	AAW93382	Fungal consensus p
15	2308	93.5	467	20	AAW93380	Fungal phytase pro
16	2308	93.5	467	21	AAW93381	Fungal consensus p
17	2308	93.5	467	21	AAW93381	Consensus phytase
18	2308	93.5	467	21	AAW93383	Phytase-1, a conse
19	2307	93.5	467	20	AAW93383	Fungal consensus p
20	2305	93.4	467	20	AAW93384	Fungal consensus p
21	2299	93.2	467	20	AAW93385	Fungal consensus p
22	2298	93.1	467	20	AAW93906	Ascomycete consens
23	2297	93.1	467	21	AAW93906	Consensus phytase
24	2297	93.1	467	21	AAW93906	Consphy12, a deriv
25	2190	88.7	467	20	AAW93906	Consensus phytase-
26	2190	88.7	467	20	AAW93906	Phytase-7, a deriv
27	2187	88.6	467	21	AAW93906	Consensus phytase
28	2182	88.4	441	21	AAW93906	Consensus phytase
29	2093	84.8	437	21	AAW93906	Consensus phytase
30	2093	84.8	437	21	AAW93906	Phytase-11, a cons
31	2056	83.3	431	21	AAW93906	Consensus phytase
32	2056	83.3	431	21	AAW93906	Initial consensus
33	2046	82.9	467	21	AAW93906	Mutant Aspergillus
34	2028.5	82.2	424	21	AAW93906	Consensus phytase
35	2028.5	82.2	424	21	AAW93906	Initial consensus
36	2026	82.1	467	20	AAW93906	A. fumigatus phyta
37	2026	82.1	467	21	AAW93906	Aspergillus fumiga
38	2026	82.1	467	21	AAW93906	Mutant Aspergillus
39	1888	76.5	465	19	AAW93906	Aspergillus fumiga
40	1888	76.5	465	20	AAW93906	A. fumigatus phyta
41	1878.5	76.1	440	21	AAW93906	Aspergillus fumiga
42	1875.5	76.0	440	21	AAW93906	Aspergillus fumiga
43	1875.5	76.0	440	21	AAW93906	Aspergillus fumiga
44	1870.5	75.8	440	21	AAW93906	Aspergillus fumiga
45	1870.5	75.8	440	21	AAW93906	Aspergillus fumiga

ALIGNMENTS

RESULT 1

AA43170

ID AA43170 standard; Protein; 467 AA.

XX

AC AA43170;

XX

DT 06-JAN-2000 (first entry)

XX

XX Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.

DE

XX

KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;

KW consensus sequence.

XX

OS Synthetic.

XX

PN WO9948380-A1.

XX

PD 30-SEP-1999.

XX

PF 22-MAR-1999; 99WO-DK00154.

XX

PR 23-MAR-1998; 98DK-0000407.

PR 19-JUN-1998; 98DK-0000806.

PR 18-SEP-1998; 98DK-0001176.

PR 22-JAN-1999; 99DK-0000091.

XX

XX 22-JAN-1999; 99DK-0000093.

PA (NOVO) NOVO-NORDISK AS.

XX

PI Petersen S;

XX

DR WPI; 1999-591030/50.

DR N-PSDB; AA231521.

XX

PT Preparing animal feed using a thermostable phytase -

XX Example 3; Fig 10; 71pp; English.

XX This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 20; Length 467;
 PS Best Local Similarity 100.0%; Pred. No. 2.6e-237;
 XX Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDTVDGGYQCPEISHLWGTYSPPFSLADE 60
 DB 1 mgvfvlslatlfgstgtaglprgnshscdtvdggyqcpeishlwtgyspfslade 60
 QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 saispdvpkgrvtvfqvlslrhgaryptssaskaysalieaiqknatafkgyaflktn 120
 QY 121 YTLGADDLTPFGEQMNVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGFSQA 180
 DB 121 ytlgaddltpfgeqmnvnsqkifryrykalkarkivpfirasgsdrviasaekflegfsa 180
 QY 181 KLADPGANPHOASPVINVTIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvtipegagynntldhglctafeeselgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPFCDLTFHDEWIQYDLQSLG 300
 DB 241 arleahlpvnltdedvvnldmcpfdvtartsdatsqlspfcldtfhdewiqydlqslg 300
 QY 301 KYGAGNPLGPAQGVGFVNLIELRLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfvnlielrlthspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdntmwsiffalglngtkplsttsvesieetdgyaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlndrvvplhgcgvdklgrckrdrdfveglsfarsggnweecfa 467

RESULT 2

AAB20527
 ID AAB20527 standard; Protein; 467 AA.

XX AAB20527;

DT 05-DEC-2000 (first entry)

XX Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytase degradation rate;
 XX food; feed; phytate; manure.

OS Synthetic.

XX WO200043503-A1.

PN

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX N-PSDB; AAA73234.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 4; Fig 8a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytase degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 21; Length 467;
 PS Best Local Similarity 100.0%; Pred. No. 2.6e-237;
 XX Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDTVDGGYQCPEISHLWGTYSPPFSLADE 60
 DB 1 mgvfvlslatlfgstgtaglprgnshscdtvdggyqcpeishlwtgyspfslade 60
 QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 saispdvpkgrvtvfqvlslrhgaryptssaskaysalieaiqknatafkgyaflktn 120
 QY 121 YTLGADDLTPFGEQMNVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGFSQA 180
 DB 121 ytlgaddltpfgeqmnvnsqkifryrykalkarkivpfirasgsdrviasaekflegfsa 180
 QY 181 KLADPGANPHOASPVINVTIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvtipegagynntldhglctafeeselgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPFCDLTFHDEWIQYDLQSLG 300
 DB 241 arleahlpvnltdedvvnldmcpfdvtartsdatsqlspfcldtfhdewiqydlqslg 300
 QY 301 KYGAGNPLGPAQGVGFVNLIELRLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfvnlielrlthspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdntmwsiffalglngtkplsttsvesieetdgyaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlndrvvplhgcgvdklgrckrdrdfveglsfarsggnweecfa 467

RESULT 3

AAV69569
ID AAV69569 standard; Protein; 467 AA.

XX
AC AAV69569;

DT 19-APR-2000 (first entry)

DE Mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
KW thermostable; animal feed; monogastric animal; phytate phosphorus;
KW phosphate availability; consensus; mutant; mutein.

XX OS Aspergillus terreus 9A1

OS OS Aspergillus terreus cbs16.46.

OS OS Aspergillus niger var. awamori.

OS OS Aspergillus niger str. NRRL3135.

OS OS Aspergillus fumigatus ATCC13073.

OS OS Aspergillus fumigatus ATCC32722.

OS OS Aspergillus fumigatus ATCC58128.

OS OS Aspergillus fumigatus ATCC26906.

OS OS Aspergillus fumigatus ATCC32239.

OS Emericella nidulans.

OS Talaromyces thermophilus ATCC20186.

OS Myceliophthora thermophila.

OS Paxillus involutus NN005693.

OS Trametes pubescens NN9343.

OS Agrocybe pediades NN009289.

OS Peniophora lycii NN006113.

OS Thermomyces lanuginosa.

OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /note= "Phytase signal peptide from Aspergillus terreus

FT Protein 27..467

FT /note= "Mature phytase-10-thermo[3]-Q50T-K91A"

XX EP969089-A1.

XX 05-JAN-2000.

XX 23-JUN-1999; 99EP-0111949.

XX 29-JUN-1998; 98EP-0111960.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Bruggen R, Lehmann M, Wyss M;

XX WPI; 2000-099429/09.

XX New stabilised enzyme formulation, useful for feed compositions for

XX monogastric animals

XX Example 5; Fig 19; 101pp; English.

XX The invention relates to a novel stabilised dry or liquid enzyme

XX formulation, comprising phytase (myo-inositol hexakisphosphate

XX phosphohydrolase) and one or more stabilising agents including

XX xylitol or ribitol; polyethylene glycols with a molecular weight of 600

XX to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,

XX glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.

XX The stabilised phytase formulation is used in a method for preparing a

XX feed composition for monogastric animals (e.g., pigs, poultry) and

XX provides a monogastric animal with its dietary requirements of

XX phosphorus. Although a large amount of phosphate is present in animal

XX feed in the form of phytate phosphorus, monogastric animals are unable

XX to utilise this form of phosphate, resulting in the addition of extra

CC phosphate to the feed of such animals. Phytase enhances the nutritional
CC value of plant material without the need for adding additional phosphate
CC to the feed. The level of phosphate pollution in the environment is
CC reduced by adding phytase to animal feed, as the animal can make use of
CC the inorganic phosphate liberated from phytate phosphorus using the
CC enzyme. The phytase formulation of the invention has an improved
CC thermostability and can therefore remain stable during long-term storage
CC and can withstand feed processing methods such as extrusion, expansion
CC and pelleting. The present sequence represents a mutant phytase-10
CC consensus sequence, phytase-10-thermo[3]-Q50T-K91A, which has a
CC temperature optimum and melting point 4 degrees Celsius higher than that
CC of phytase-10 (AAV69566). Its specific activity with phytate as a
CC substrate is also strongly increased.

XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 21; Length 467;

Best Local Similarity 100.0%; Pred. No. 2.6e-237;

Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGCTALGPRGNSHSDTVGGYOCFPEISHLWMTYSPFFSLADE 60

DB 1 mgvfvlslatlfgstgctalgrgnshtcdtdggycfpeishlwtyspffslade 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAFKGYAFLLKTYN 120

DB 61 saispdvpgkcrvtfvqvlshrhgaryptssaskaysalieleaigknatarfkyafllkty 120

QY 121 YTLGADLTTPFGEQOMVNSGIKFRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180

DB 121 ytlgaddltfpgeqomvnsigikfrykalarikiwpfirasgsdrviasaeakfiqfqa 180

QY 181 KLADGAPNHPQASPVINVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240

DB 181 kladgaanphqaspvinvipegagynntldhglctafeeselgddveanftavfappir 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSDFCDLTFHDEWIOYDIQSLG 300

DB 241 arleahlpgvnltdedvnlmdmcpfdtvtartsdatsdqspfdcltfhdewidgydyiqslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

DB 301 kyygygagnplgpaqgvgfvneliarlthspvqdhstnthtldsnpatplnatlyadfs 360

QY 361 HDNTWVSIFPFFALGLYNGTKPLSTTSVESTETEDGYSASWTVPFAARAYVEMMQCEAKEP 420

DB 361 hdntmvsiffalglyngtkplsttsvesietdgyaswtvpfaarayvemmqceakep 420

QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDPVEGLSFARSGGNWEEGFA 467

DB 421 lvrvlvndrvvplhgcvdklgcrdrdpveglstarsggnweecfa 467

RESULT 4

AAAB20534

ID AAB20534 standard; Protein; 467 AA.

XX AC AAB20534;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;

XX temperature stability; pH profile; temperature profile; reaction rate;

XX specific activity; substrate specificity; substrate cleavage pattern;

XX substrate binding; position specificity; phytate degradation rate;

XX food; feed; phytate; manure.

XX OS Synthetic.

XX WO200043503-A1.

XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-DK00025.
XX PR 22-JAN-1999; 99DK-0000092.
XX PR 21-SEP-1999; 99DK-0001340.
XX PA (NOVO) NOVO NORDISK AS.
XX PI Lehmann M;
XX DR WPI; 2000-491161/43.
XX DR N-PSDB; AAA73293.
XX PT Novel phytases with improved properties such as temperature stability,
XX PT pH stability and substrate specificity, for use in pharmaceuticals and
XX PT compound foods and feeds -
XX PS Disclosure; Fig 25a-c; 240pp; English.
XX CC The present invention describes improved phytases, preferably with
XX CC increased thermostability, and methods for producing them. The methods
XX CC can be used for producing phytases with improved properties e.g.
XX CC temperature stability, pH stability, pH profile, temperature profile,
XX CC specific activity, substrate specificity, substrate cleavage pattern,
XX CC substrate binding, position specificity, the velocity and level of
XX CC release of phosphate from corn, reaction rate, phytate degradation rate,
XX CC and end level of released phosphate. The phytases can be used to produce
XX CC pharmaceutical compositions or compound food or feeds. The feed can be
XX CC used to reduce levels of phytate in animal manure, by converting it
XX CC into lower inositol phosphates and/or inositol and inorganic phosphate.
XX CC The present sequence represents a phytase sequence from the present
XX CC invention.
XX SQ Sequence 467 AA;
Query Match 99.5%; Score 2456; DB 21; Length 467;
Best Local Similarity 99.6%; Pred. No. 4.1e-236;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MGFFVLLSIATLFGSTGTALGRGNHSCDVTDDGGYQCFFPEISHLWGTYSPPFFSLADE 60
DB 1 mgffvllsiatlfgstgtalgrgnshscdvtddggycfpeishlwgtyspffslade 60
QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFAKGYAFLKTYN 120
DB 61 saispdvpkgcrvtfvqlsrhgaryptssaskaysalieaiqknafakgyaflkty 120
QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 ytlgaddltpfgqomvnschkfryrykalkarkivpfiirsgsdrrviasaekfiqf 180
QY 181 KIADPGANPHQASPVNIIIEGAGYNNLTHGLCTAFEESELGDDVEANTVAFAPPIR 240
DB 181 kiadpganphqaspvniiliegagynnlthglctafeeselgddveantvafappir 240
QY 241 ARLEAHLPGVNLTDDEYVNLMDMCPEDTAVRTSDATQLSPFCDLTFHDEWIOYDLOSLG 300
DB 241 arleahlpgvnltdedvynlmdmcpedtavrtsdatlspfcldtfhdewioydylosl 300
QY 301 KYGYGAGNPLGPAGGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kygygagnplgpaggvfneliarlthspvqdhtstnhtldsnpatfplnatlyadfs 360
QY 361 HNTWVSIFFAFGALYNGKPKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCEAKEP 420
DB 361 hntwvsiffalngkpkplsttsvesieetdgyasaswtvpfaarayvemmqceakep 420
QY 421 LVRLVNDRVVPLHCGGVKDLGRCKRDDFVEGLGFARSGGNWEECFA 467
DB 421 lvrlvndrvvplhcggvkdlgrckrddfveglgfarsgggnweecfa 467

RESULT 5
AAB20533
ID AAB20533 standard; Protein; 467 AA.
XX AC AAB20533;
XX DT 05-DEC-2000 (first entry)
XX DE Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95.
XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
XX KW temperature stability; pH profile; temperature profile; reaction rate;
XX KW specific activity; substrate specificity; substrate cleavage pattern;
XX KW substrate binding; position specificity; phytate degradation rate;
XX KW food; feed; phytate; manure.
XX OS Synthetic.
XX PN WO2000043503-A1.
XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-DK00025.
XX PR 22-JAN-1999; 99DK-0000092.
XX PR 21-SEP-1999; 99DK-0001340.
XX PA (NOVO) NOVO NORDISK AS.
XX PI Lehmann M;
XX DR WPI; 2000-491161/43.
XX DR N-PSDB; AAA73292.
XX PT Novel phytases with improved properties such as temperature stability,
XX PT pH stability and substrate specificity, for use in pharmaceuticals and
XX PT compound foods and feeds -
XX PS Disclosure; Fig 24a-c; 240pp; English.
XX CC The present invention describes improved phytases, preferably with
XX CC increased thermostability, and methods for producing them. The methods
XX CC can be used for producing phytases with improved properties e.g.
XX CC temperature stability, pH stability, pH profile, temperature profile,
XX CC specific activity, substrate specificity, substrate cleavage pattern,
XX CC substrate binding, position specificity, the velocity and level of
XX CC release of phosphate from corn, reaction rate, phytate degradation rate,
XX CC and end level of released phosphate. The phytases can be used to produce
XX CC pharmaceutical compositions or compound food or feeds. The feed can be
XX CC used to reduce levels of phytate in animal manure, by converting it
XX CC into lower inositol phosphates and/or inositol and inorganic phosphate.
XX CC The present sequence represents a phytase sequence from the present
XX CC invention.
XX SQ Sequence 467 AA;
Query Match 99.3%; Score 2451; DB 21; Length 467;
Best Local Similarity 99.4%; Pred. No. 1.3e-235;
Matches 464; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MGFFVLLSIATLFGSTGTALGRGNHSCDVTDDGGYQCFFPEISHLWGTYSPPFFSLADE 60
DB 1 mgffvllsiatlfgstgtalgrgnshscdvtddggycfpeishlwgtyspffslade 60
QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNAFAKGYAFLKTYN 120
DB 61 saispdvpkgcrvtfvqlsrhgaryptssaskaysalieaiqknafakgyaflkty 120
QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 ytlgaddltpfgqomvnschkfryrykalkarkivpfiirsgsdrrviasaekfiqf 180

Db 121 ytlgaddltpfgeqgmnsigikfyrkalkarivfirasgdrviaseakfiegfsa 180
 QY 181 KLADGANPHQASPVINVIPEAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 Db 181 kladganphqaspviniipeagaynnitldhglctafeestlgddveanftavfapp 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSDFTHDEWIOYDYLQSLG 300
 Db 241 arleahlpvnltdedvnlmdmcpdftvartsdatsqlspfcldfthdewigydyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 Db 301 kygygagnplgpagvgvneliarlthspvgdhtstnhtldsnpatplnatlyadfs 360
 QY 361 HDNTMVSIFPFGALGNGTKPLSTTSVESTTEEDGYASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntmvsiffalglyngtkplsttsvestteedgyaswvcpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrdrdfveglstfarsggnweecfa 467

RESULT 6

AAB20524
 ID AAB20524 standard; Protein: 467 AA.

XX AC AAB20524;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 10 SEQ ID NO:26.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 XX food; feed; phytate; manure.

OS Synthetic.

XX WO200043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK000025.

XX PR 22-JAN-1999; 99DR-00000092.

XX PR 21-SEP-1999; 99DR-0001340.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX WPI: 2000-491161/43.

XX N-PSDB; AAA73232.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 1; Fig 5a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g. file,
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it

CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 99.2%; Score 2448; DB 21; Length 467;

Best Local Similarity 98.9%; Pred. No. 2.6e-235;

Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNHSCDVGQYQCFPEISHLWGTYSFFSLADE 60

Db 1 mgvfvllsiatlfgstgtalgrgnshscdvdggyqcfpeishlwgyqsfslade 60

QY 61 SAISDPVPRGCRVTFVQVLSRHGARYPTSSAKAYSALIEATQKATAFKGVAFUKTYN 120

Db 61 saispdvprgcrvtfvqlsrhgaryptsakkskysallieaigknatafkgvafukty 120

QY 121 YTLGADDLTPFGEQGMNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180

Db 121 ytlgaddltpfgeqgmnsigikfyrkalkarivfirasgdrviaseakfiegfsa 180

QY 181 KLADGANPHQASPVINVIPEAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240

Db 181 kladganphqaspviniipeagaynnitldhglctafeestlgddveanftavfapp 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSDFTHDEWIOYDYLQSLG 300

Db 241 arleahlpvnltdedvnlmdmcpdftvartsdatsqlspfcldfthdewigydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

Db 301 kygygagnplgpagvgvneliarlthspvgdhtstnhtldsnpatplnatlyadfs 360

QY 361 HDNTMVSIFPFGALGNGTKPLSTTSVESTTEEDGYASWTVPFAARAYVEMMQCEAKEP 420

Db 361 hdntmvsiffalglyngtkplsttsvestteedgyaswvcpfaarayvemmqceakep 420

QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467

Db 421 lvrvlndrvplhgcavdklgrckrdrdfveglstfarsggnweecfa 467

RESULT 7

XX ID AAY69566 standard; protein: 467 AA.

XX AC AAY69566;

XX DT 19-APR-2000 (first entry)

XX DE Phytase-10, a consensus phytase.

XX KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;

XX KW thermostable; animal feed; monogastric animal; phytate phosphorus;

XX KW phosphate availability; consensus; phytase-10.

XX OS Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS Aspergillus niger var. awamori.

OS Aspergillus niger str. NRRL3135.

OS Aspergillus fumigatus ATCC13073.

OS Aspergillus fumigatus ATCC32722.

OS Aspergillus fumigatus ATCC58128.

OS Aspergillus fumigatus ATCC26906.

OS Emericella nidulans.

OS Talaromyces thermophilus ATCC20186.

OS Myceliophthora thermophila.

OS Paxillus involutus NN005693.

OS Trametes pubescens NN9343.

OS Agrocybe pediades NN009289.

OS Peniophora lycii N0006113.
OS Thermomyces lanuginosa.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..26
XX FT /note= "Phytase signal peptide from Aspergillus terreus
XX FT cbs16.46"
XX FT Protein 27..467
XX FT /note= "Mature phytase-10 consensus"
XX
XX EP960899-Al
XX PD 05-JAN-2000.
XX
XX 23-JUN-1999; 99EP-0111949.
XX 29-JUN-1998; 98EP-0111960.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Brudger R, Lehmann M, Wyss M;
XX WPI; 2000-099429/09.
XX N-PSDB; AAZ59642.
XX
XX New stabilized enzyme formulation, useful for feed compositions for
XX monogastric animals -
XX
XX Example 4; Fig 17; 10lpp; English.
XX The invention relates to a novel stabilised dry or liquid enzyme
XX formulation, comprising phytase (myo-inositol hexakisphosphate
XX phosphohydrolase) and one or more stabilising agents including
XX xylitol or ribitol; polyethylene glycols with a molecular weight of 600
XX to 4000 Da, and preferably 1000 to 3350 Da; the disodium salts of malonic,
XX glutamic and succinic acid; carboxymethylcellulose; and sodium alginate.
XX The stabilised phytase formulation is used in a method for preparing a
XX feed composition for monogastric animals (e.g., pigs, poultry) and
XX provides a monogastric animal with its dietary requirements of
XX phosphorus. Although a large amount of phosphate is present in animal
XX feed in the form of phytate phosphorus, monogastric animals are unable
XX to utilise this form of phosphate, resulting in the addition of extra
XX phosphate to the feed of such animals. Phytase enhances the nutritional
XX value of plant material without the need for adding additional phosphate
XX to the feed. The level of phosphate pollution in the environment is
XX reduced by adding phytase to animal feed, as the animal can make use of
XX the inorganic phosphate liberated from phytate phosphorus using the
XX enzyme. The phytase formulation of the invention has an improved
XX thermostability and can therefore remain stable during long-term storage
XX and can withstand feed processing methods such as extrusion, expansion
XX and pelleting. The present sequence represents a consensus sequence,
XX designated phytase-10, which was derived from the mature phytase
XX sequences from a variety of fungi (AAV69544-Y69546, AAV69548-Y69556,
XX AAV69564) and the Basidiomycetes phytase consensus AAV69563 and
XX additionally contains the Aspergillus terreus cbs116.46 signal peptide at
XX the N-terminus.
XX
XX Sequence 467 AA;
XX SQ

Query Match 98.2%; Score 2448; DB 21; Length 467;
Best Local Similarity 98.9%; Pred. No. 2.6e-235;
Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVPVLLSTATLGSTGALPRGNHSCDVTGCGYOCFFPEISHLWCTYSPFFSLADE 60
DB 1 mgvfvllsialtfgstgslgprgnshscdvtvggycgfpfshlwggyspfislaide 60
QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAPKGYAFILKTYN 120
DB 61 saispdvpkgrvtfvqvlshrhyrptsskksyallieaiaqknatarkgyafilktyn 120

QY 121 YTLGADDLTPFGGEQQMVNSGKIFRYRYKALARKIYVFFIRASGSDRVIASAEKFIQFOSA 180
DB 121 ytlgaddltpfgeqqmvnsgikfyrryrykalkarkivpfvrasgsdrviasaekfiqfosa 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAPFESSELGDDVEANFTAVFAPPIR 240
DB 181 kladpganphqaspvinvipegagynntldhglctapeeselselgddveanftavfappir 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCMCPDVTARTSDATQLSPPCDLFTHDEWIQYDYLQSLG 300
DB 241 arleahlpgvnltddevnvlmdcmcpdvtartsdatsqlspfcldlthdewiqydyqlslg 300
QY 301 KYCYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPEPLNATLYADFS 360
DB 301 kycygagnplgpaqgvgfvneliarlthspvqgdhtstnhtlidsnpatfplnatlyadfs 360
QY 361 HDNTMVSIFPAGLYNGTKPLSTTSVESTTEEDGYSASWTVPFAARAYVEMMOCAEKEP 420
DB 361 hdntmvsiffalglyngtkplsttsvestteedgyaswtvpfaarayvemmqceakep 420
QY 421 LVRVLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSNGNNEECFA 467
DB 421 lvrvlvndrvvplhggvdklgrckrddfvglslfarsngnneecfa 467
RESULT 8
AAB20532
ID AAB20532 standard; Protein; 467 AA.
XX AC AAB20532;
XX DT 05-DEC-2000 (first entry)
XX DE Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.
XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
XX KW temperature stability; pH profile; temperature profile; reaction rate;
XX KW specific activity; substrate specificity; substrate cleavage pattern;
XX KW substrate binding; position specificity; phytate degradation rate;
XX KW food; feed; phytate; manure.
XX OS Synthetic.
XX WO2000043503-A1.
XX PD 27-JUL-2000;
XX 21-JAN-2000; 2000WO-DK00025.
XX 22-JAN-1999; 99DK-0000092.
XX 21-SEP-1999; 99DK-0001340.
XX (NOVO) NOVO NORDISK AS.
XX Lehmann M;
XX WPI; 2000-491161/43.
XX N-PSDB; AAA73291.
XX Novel phytases with improved properties such as temperature stability,
XX pH stability and substrate specificity, for use in pharmaceuticals and
XX compound foods and feeds -
XX Disclosure; Fig 23a-c; 240pp; English.
XX The present invention describes improved phytases, preferably with
XX increased thermostability, and methods for producing them. The methods
XX can be used for producing phytases with improved properties e.g.
XX temperature stability, pH stability, pH profile, temperature profile,
XX specific activity, substrate specificity, substrate cleavage pattern,
XX substrate binding, position specificity, the velocity and level of
XX release of phosphate from corn, reaction rate, phytate degradation rate,
XX and end level of released phosphate. The phytases can be used to produce

CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX
 SQ Sequence 467 AA;

Query Match 95.9%; Score 2367; DB 21; Length 467;
 Best Local Similarity 95.7%; Pred. No. 3.1e-227;
 Matches 447; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGQYCFPEISHLWGTSPFFSLADE 60
 DB 1 mgvfvlslatlfgstgaltalpgnshscdvtgqyqcfpeishlwgtspffslade 60

QY 61 SAISDPVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAPKGYAFILKTYN 120
 DB 61 saispdvpkdcrtvfvqlsrhgaryptssaskaysalieaoknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrrykalarkivpfirasgsdrviasaekfiqfqs 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpgsqhqsapvlnvipegsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDIQYDLSIG 300
 DB 241 arleahlpgvnltddevvnlmdcpfdvtartsdatselgddveanftalfapair 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 kygygagnpplgagvgfaneliarlthspvqdhstntldsnpatplnatlyadfs 360

QY 361 HNTMVSIFPAGLNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hntmvsifpaglgyngtkplsttsvesieetdgyaswtvpfaarayvemmqqakep 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWEECA 467
 DB 421 lvrvlndrvplhgcavdklgrckdrdfveglsfarsggnwaeca 467

RESULT 9

AA43169
 ID AA43169 standard; Protein; 467 AA.

XX
 AC AA43169;

XX
 DT 06-JAN-2000 (first entry)

XX
 DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.

XX
 KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 consensus sequence.

XX
 OS Synthetic.

XX
 PN WO9948380-A1.

XX
 PD 30-SEP-1999.

XX
 PF 22-MAR-1999; 99WO-DK00154.

XX
 PR 23-MAR-1998; 98DK-0000407.

XX
 PR 19-JUN-1998; 98DK-0000806.

XX
 PR 18-SEP-1998; 98DK-0001176.

XX
 PR 22-JAN-1999; 99DK-0000091.

XX
 PR 22-JAN-1999; 99DK-0000093.

PA (NOVO) NOVO-NORDISK AS.

XX
 PI Petersen S;

XX
 DR WPI; 1999-591030/50.
 N-PSDB; AA231520.

XX
 PT Preparing animal feed using a thermostable phytase

XX
 XX Example 3; Fig 9; 7lpp; English.

XX
 CC This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX
 SQ Sequence 467 AA;

Query Match 95.7%; Score 2362; DB 20; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGQYCFPEISHLWGTSPFFSLADE 60
 DB 1 mgvfvlslatlfgstgaltalpgnshscdvtgqyqcfpeishlwgtspffslade 60

QY 61 SAISDPVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAPKGYAFILKTYN 120
 DB 61 saispdvpkdcrtvfvqlsrhgaryptssaskaysalieaoknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrrykalarkivpfirasgsdrviasaekfiqfqs 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpgsqhqsapvlnvipegsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDIQYDLSIG 300
 DB 241 arleahlpgvnltddevvnlmdcpfdvtartsdatselgddveanftalfapair 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 kygygagnpplgagvgfaneliarlthspvqdhstntldsnpatplnatlyadfs 360

QY 361 HNTMVSIFPAGLNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hntmvsifpaglgyngtkplsttsvesieetdgyaswtvpfaarayvemmqqakep 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWEECA 467
 DB 421 lvrvlndrvplhgcavdklgrckdrdfveglsfarsggnwaeca 467

RESULT 10

AA20526

ID AA20526 standard; Protein; 467 AA.

XX
 AC AA20526;

XX
 DT 05-DEC-2000 (first entry)

XX
 DE Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 temperature stability; pH profile; temperature profile; reaction rate;

KW specific activity; substrate specificity; phytase; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA:

Query Match 95.7%; Score 2362; DB 21; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCFPEISHLWGTYSPPFFSLADE 60
 DB 1 mgffvllsiatlfgstgtalgpgrnshscdtpdggycfpelshlwtgyspyfslade 60
 QY 61 SAISDPVPGKCRVTFVQLSRGARYPTSSASKAYSALIEAQKATAFKGYAFLKTYN 120
 DB 61 saisdvpdkcrvtfvqlsrgraryptssaskaysalieaqlkntafkgyafktycn 120
 QY 121 YTLGADDLTPFGEOQMNWNGIKFYRYKALARKIYPTFRASGSDRVIASAEKFTIGFQSA 180
 DB 121 ytlgaddltpfgeqomnwngikfyrkylarkiyptfrasgsdrviasaekftigfqs 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNNTLDHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
 DB 181 kladpgsqhqpaspvinviipegagynntldhgtctafedseelgddveanftalfapair 240
 QY 241 ARLEAHLPGVNLTDENVLMDCPFDTVARTSDATOLSPFCDLTHDEWQYDYLQSLG 300
 DB 241 arleahlpvnltdenvlmcpfdtvtartsdatslpcfdlthdewqdydylqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSNHTLDSNPATFPFNATLYADFS 360
 DB 301 kyygygagnpapgvgvfanellarlthspvqdhstnhtldsnpatfpfnatlyadfs 360

QY 361 HDNTMVSIFALGLYNGTKPLSTSTSVESIEETDGYSSASWTVPFAARAYVEMQCEAKEP 420
 DB 361 hdnmtvisiffalglyngtkplststsvesieetdgyssaswtvpfaarayvemmqcqeakep 420
 QY 421 LVRVLYNDRVVPVPHGCGVDKLGRCRDRDDFVEGLSFARSGNWECEFA 467
 DB 421 lvrvlyndrvvpvphgcvdklgrckrddrddfveglsfarsgnwaecefa 467

RESULT 11

AAB20531

ID AAB20531 standard; Protein; 467 AA.

XX AAB20531;

XX 05-DEC-2000 (first entry)

DE Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

XX WO2000043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.7%; Score 2362; DB 21; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCFPEISHLWGTYSPPFFSLADE 60
 DB 1 mgffvllsiatlfgstgtalgpgrnshscdtpdggycfpelshlwtgyspyfslade 60

QY 421 LVRVLNDRVPLHGGGVDKLGRCKRDDFVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlndrvvplhgcavdkgkrckrddfvglsgfarsggnwaecfa 467

RESULT 13

AAAB20523
 ID AAB20523 standard; Protein; 441 AA.

AC AAB20523;

XX 05-DEC-2000 (first entry)

DE Consensus phytase 10 (Fcp10) SEQ ID NO:24.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Example 2; Fig 4a-d; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 441 AA;

Query Match 94.1%; Score 2322; DB 21; Length 441;
 Best Local Similarity 98.9%; Pred. No. 8.6e-223;
 Matches 436; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 27 NSHSCDFVGGYQCFPIHSLWGTYSPPFFSLADESAISPDPVPGKCRVTFVQLSRGARY 86
 DB 1 nshscdfvgyqcfpishlwgtysppffsladesaisdpvpgkcrvtivqlsrngary 60

QY 87 PTSSASRAYSALTEATOKNATKGYAFKTYNTLGDADLTPFGQQMVNSGKIFYRR 146
 DB 61 ptsskskysallteatqknatfkgayfklktyntlgadldtpfgeqmvnsqikfyr 120

QY 147 YKALARKIVPFRASGSDRVIASAEKFIKGFOSAKLADPGANPHOASPVINVIPEGAGY 206
 DB 121 ykalararkivpfrasgsdrviasaekflegfsakladpganphqspvinvilpegagy 180
 QY 207 NNTLDHGLCTAFEESELGDDEANFTAVFAPPFIRARLEAHLPGVNLTDEDVYNLMDMCPF 266
 DB 181 nntldhglctafeeseldgddveanftavfappfrrarleahlpgvnltdedvynlmdmcpf 240
 QY 267 DTVARTSDATQSLSPFCDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAGCGFVNELIAR 326
 DB 241 dtvartsdattqlspfcldtfhdewiodylqslgkygygagngpaggvgnvneliar 300
 QY 327 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTWVSIFFALGLYNGTKPLSTTSV 386
 DB 301 lthspvqdhstnthtldsnpatplnatlyadfsdhtnwmvsiffalglyngtkplsttsv 360
 QY 387 ESIEETDGYASWTVPFAARAYVEMMOCEAEKPLVRVLNDRVPLHGGGVDKLGRCKR 446
 DB 361 esieetdgyaaswtvpfaarayvemmocaekeplvrvlndrvvplhgcvdklgrckr 420
 QY 447 DDFVEGLSFARSGGNWEECFA 467
 DB 421 ddfveglsfarsggnweecfa 441

RESULT 14

AAW93382
 ID AAW93382 standard; Protein; 467 AA.

XX AAW93382;

XX 11-JUN-1999 (first entry)

XX Fungal consensus phytase protein mutant Q50T.

XX Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; mutein; feed; pharmaceutical.

XX Fungi.

XX Synthetic.

XX PN EP897985-A2.

XX PD 24-FEB-1999.

XX PF 15-JUL-1998; 98EP-0113176.

XX PR 24-JUL-1997; 97EP-0112688.

XX PA (HOFF) HOFFMANN LA ROCHE AG F.

XX Lehmann M;

XX WPI; 1999-134647/12.

XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences

XX Claim 9; Page -; 30pp; English.

XX This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or mutein. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and mutein are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.

XX Sequence 467 AA;

Query Match 93.8%; Score 2314; DB 20; Length 467;
 Best Local Similarity 93.6%; Pred. No. 5,9e-222;
 Matches 437; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGRGNHSCDTPDGGYQCFPEISHLWGTSPFFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshscdtpdggycfpeishlwtgtsyfyfslade 60

QY 61 SAISPDVPKCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKATAFKGYAFLKTYN 120
 Db 61 saispvdpdcvrtfvqlsrhgarypteskskaysalieatqknatafkgyafkty 120

QY 121 YTLGADDLTPFGQQMVNSGIKFYRRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 Db 121 ytlgaddltpfgengmvsngikfyrkalarakiyfirasgsdrviasaeakfiqfqs 180

QY 181 KLADPCANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpcahpqaspvlnviipegagynntldhglctafeslgeddveanftalfap 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPEDTVARTSDATQSPFCDLTFHDEWIOYDYLQSLG 300
 Db 241 arleahlpgvntldedvnlmdmcpetvartsdatspfcldtfhdewrydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTMVSIFFAALGYNKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntmsiffaalgynkplsttsvesieetdgyaswtvpfagarayvemmqc 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrddrfveglsfarsggnwaecfa 467

RESULT 15

AAW93380
 ID AAW93380 standard; Protein; 467 AA.
 AC AAW93380;
 DT 11-JUN-1999 (first entry)
 DE Fungal phytase protein consensus DNA.
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 OS feed additive; variant; mutein; feed; pharmaceutical.
 OS Fungi.
 OS Synthetic.
 PN EP997985-A2.
 PD 24-FEB-1999.
 PF 15-JUL-1998; 98EP-0113176.
 PR 24-JUL-1997; 97EP-0112688.
 PR (HOFF) HOFFMANN LA ROCHE AG F.
 PA Lehmann M;
 PI WPI; 1999-134647/12.
 DR N-PSDB; AAX23022.
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 XX Claim 8; Fig 2; 30pp; English.
 PS

XX This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses
 CC phytase to valuable feed additives, with a fully defined amino acid
 CC sequence given in the specification, or variant or mutant. The method is
 CC useful for improving protein properties by altering their sequence. The
 CC consensus protein and mutein are useful in food, feed or pharmaceutical
 CC compositions. This sequence represents the consensus phytase protein used
 CC in the method of the invention.
 XX
 SQ Sequence 467 AA;

Query Match 93.5%; Score 2308; DB 20; Length 467;
 Best Local Similarity 93.4%; Pred. No. 2.3e-221;
 Matches 436; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGRGNHSCDTPDGGYQCFPEISHLWGTSPFFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshscdtpdggycfpeishlwtgtsyfyfslade 60

QY 61 SAISPDVPKCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKATAFKGYAFLKTYN 120
 Db 61 saispvdpdcvrtfvqlsrhgarypteskskaysalieatqknatafkgyafkty 120

QY 121 YTLGADDLTPFGQQMVNSGIKFYRRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 Db 121 ytlgaddltpfgengmvsngikfyrkalarakiyfirasgsdrviasaeakfiqfqs 180

QY 181 KLADPCANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpcahpqaspvlnviipegagynntldhglctafeslgeddveanftalfap 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPEDTVARTSDATQSPFCDLTFHDEWIOYDYLQSLG 300
 Db 241 arleahlpgvntldedvnlmdmcpetvartsdatspfcldtfhdewrydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTMVSIFFAALGYNKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntmsiffaalgynkplsttsvesieetdgyaswtvpfagarayvemmqc 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrddrfveglsfarsggnwaecfa 467

Search completed: October 26, 2001, 16:40:05
 Job time: 4956 sec

us-09-488-265-31.rag

Sat Oct 27 15:25:53 2001

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:47 ; Search time 32.57 seconds
(without alignments)
295.232 Million cell updates/sec

Title: US-09-488-265-31
Perfect score: 2468
Sequence: 1 MGFFVLLSIATLFGSTGT.....DFVEGLSFARSGNNECF 467

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2182	88.4	441	4	US-09-121-425-1
2	2170	87.9	467	4	US-09-121-425-2
3	1852	75.0	467	1	US-07-923-724-8
4	1852	75.0	467	2	US-08-609-426A-8
5	1852	75.0	467	2	US-08-374-652C-2
6	1848	74.9	467	1	US-08-151-574-32
7	1848	74.9	467	1	US-08-146-424-20
8	1848	74.9	467	1	US-08-693-709-2
9	1848	74.9	467	2	US-08-419-448-32
10	1848	74.9	467	2	US-08-819-825-3
11	1848	74.9	467	4	US-09-163-642-3
12	1841	74.6	467	4	US-09-155-855-3
13	1799	72.9	443	4	US-09-155-855-1
14	1794	72.7	443	4	US-09-155-855-2
15	1381	56.0	475	2	US-08-819-825-2
16	1381	56.0	475	4	US-09-163-642-2
17	845.5	34.3	443	3	US-08-993-359-30
18	833.5	33.8	443	3	US-08-993-359-22
19	830.5	33.7	439	3	US-08-993-359-24
20	830.5	33.7	439	3	US-09-221-654-2
21	830.5	33.7	439	3	US-08-989-358A-2
22	825	33.4	442	3	US-08-993-359-28
23	814	33.0	442	3	US-08-993-359-26
24	359	14.5	468	1	US-07-627-539G-2
25	354.5	14.4	446	1	US-07-627-539G-7
26	354.5	14.4	479	1	US-07-923-724-2
27	354.5	14.4	479	2	US-08-609-426A-2

28 354.5 14.4 479 2 US-08-374-652C-4 Sequence 4, Appli
29 185.5 7.5 92 3 US-08-993-359-32 Sequence 32, Appli
30 177.5 7.2 449 3 US-08-680-506-7 Sequence 7, Appli
31 122 4.9 113 1 US-08-241-853-8 Sequence 8, Appli
32 122 4.9 113 2 US-08-850-917-8 Sequence 8, Appli
33 120 4.9 318 3 US-08-680-506-3 Sequence 3, Appli
34 115 4.7 24 2 US-08-374-652C-32 Sequence 32, Appli
35 111 4.5 113 1 US-08-241-853-10 Sequence 10, Appli
36 111 4.5 113 2 US-08-850-917-10 Sequence 10, Appli
37 99 4.0 230 5 PCT-US93-09636-6 Sequence 6, Appli
38 99 4.0 386 1 US-08-758-213-1 Sequence 1, Appli
39 99 4.0 386 2 US-08-692-787-48 Sequence 48, Appli
40 99 4.0 386 4 US-09-097-199-48 Sequence 48, Appli
41 99 4.0 515 2 US-09-146-283-2 Sequence 2, Appli
42 99 4.0 515 3 US-08-579-823A-2 Sequence 2, Appli
43 99 4.0 515 4 US-09-344-195-2 Sequence 2, Appli
44 97 3.9 7257 3 US-09-335-409-5 Sequence 5, Appli
45 93.5 3.8 675 4 US-09-171-878-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09121,425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-1

Query Match 88.4%; Score 2182; DB 4; Length 441;
Best Local Similarity 93.0%; Pred. No. 5.7e-217;
Matches 410; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
QY 27 NSHSCDTVDGGYQCFPELSHLWGTYSPFSLADESAISPDVFKGCRVTFVQVLSRHGARY 86
Db 1 NSHSCDTVDGGYQCFPELSHLWGTYSPFSLADESAISPDVFKGCRVTFVQVLSRHGARY 60
QY 87 PTSSASKAYSALIEAIQKATAFKGYAFLKTYNTLGGADLTPEGQKVNSGIKFYRR 146
Db 61 PTSSASKAYSALIEAIQKATAFKGYAFLKTYNTLGGADLTPEGQKVNSGIKFYRR 120
QY 147 YKALARKTVPIRAGSGDRVIAIAAEKFTGFSQAKLADPGANPHQASFVINIIPGAGY 206
Db 121 YKALARKTVPIRAGSGDRVIAIAAEKFTGFSQAKLADPGANPHQASFVINIIPGAGY 180
QY 207 NNTLDHGCTAFESSEGLDDVEANFTAVFAPPIARLEAHLPGVNLNDEDDVNMDCMPF 266
Db 181 NNTLDHGCTAFESSEGLDDVEANFTAVFAPPIARLEAHLPGVNLNDEDDVNMDCMPF 240
QY 267 DTVARTSDATLSPFCDFLTHDEWIQYDIYQSLGKYGYGAGNPLGPAQGVGFNELLAR 326
Db 241 ETVARTSDATLSPFCDFLTHDEWIQYDIYQSLGKYGYGAGNPLGPAQGVGFNELLAR 300
QY 327 LTHSPVQDHTSTNITLDSNRPATFLNATLADFSHDNTWYSIFFALGLYNGTKPLSTTSV 386

Db 301 LTRSPVQDHTSTNHTLSDNPAFLPLNATLVADFSDHNSMISIFALGLYNGTAPLSTTSV 360
Qy 387 ESIEETDGYASATVPTFAARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCR 446
Db 361 ESIEETDGYASATVPTFAARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCR 420
Qy 447 DDFVEGLSFARSGGNWEECPA 467
Db 421 DDFVEGLSFARSGGNWEECPA 441

RESULT 2
US-09-121-425-2
; Sequence 2, Application 'us/09121425'
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2

Query Match 87.9%; Score 2170; DB 4; Length 467;
Best Local Similarity 85.4%; Pred. No. 1.1e-215;
Matches 416; Conservative 11; Mismatches 20; Indels 40; Gaps 2;

Qy 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCFPEISHLWGTSPFFSLADE 60
Db 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCFPEISHLWGTSPFFSLADE 60
Qy 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Qy 121 YTLGADDLTPFGEOOMVNSGIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIIEGFSQA 180
Db 101 YTLGADDLTPFGENQVNSGIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIIEGFSQA 160
Qy 181 KLADPGANPHQASPVIN-----VPIEGAGYNNVLDHGLCTAFEE 220
Db 161 KLADPGSQPHQASPVLDLIEAIQKNATAFKGYAFLKVIIEPGSGYNNVLDHGLCTAFED 220
Qy 221 SELGDDVEANFTAVFAPPTRARLEAHLPGVNTDDEVNLMDCPFDIVARTSDATQLSP 280
Db 221 SELGDDVEANFTALFAPARLEADPLGVTTDDEVVYLMDCPFEIVARTSDATELSP 280
Qy 281 FCDLFTDHEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLTHSPVDHTSTNH 340
Db 281 FCALFTDHEWRQYDYLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVDHTSTNH 340
Qy 341 TLDSNPATPLNATLVADFSDHNTWYSIFFALGLYNGTAPLSTTSVESIEETDGYASMT 400
Db 341 TLDSNPATEPLNATLVADFSDHNSMISIFALGLYNGTAPLSTTSVESIEETDGYASMT 400
Qy 401 VPFAARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCRDDFVEGLSFARSGG 460
Db 401 VPFGARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCRDDFVEGLSFARSGG 460
Qy 461 NWECEFA 467
Db 461 NWECEFA 467

Db 461 NWECEFA 467

RESULT 3
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA: US 07/044, 077
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8

Query Match 75.0%; Score 1852; DB 1; Length 467;
Best Local Similarity 74.3%; Pred. No. 7.9e-183;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCFPEISHLWGTSPFFSLADE 60
Db 1 MGVSALLPLLVLLAGVTSGLVAPASRNQSCDTPDGGYQCFSETSHLWGTSPFFSLANE 60
Qy 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120

```

QY 121 YTLGADDLTPFGQGMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGQGLVNSGKIFRYQYRESLTNRNIPFIASGSSRVIASGERFIEGFQST 180
QY 181 KLADPCANPHQASPVNIIPGAGYNWTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPCANPHQASPVNIIPGAGYNWTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCMCPDVTARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
Db 241 ARLEAHLPGVNLTDDEVNLMDCMCPDVTARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
QY 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
Db 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
QY 361 HNTMVSIFFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAVEMMOCEAEKEP 420
Db 361 HNTMVSIFFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAVEMMOCEAEKEP 420
QY 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECFA 467
Db 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECFA 467

```

RESULT 4

```

US-08-609-426A-8
; Sequence 8, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torckeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.

```

```

; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-426A-8

```

```

Query Match 75.0%; Score 1852; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 7.9e-183;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

```

```

QY 1 MGVFVLLSTATLFGSTGTALPGRNHSCDVTVDGQYCFPEISHLWGTYSFFFLADE 60
Db 1 MGVSAVLLPLLYLAGVTSLGAVPASRNQSTCDTVDGQYCFSETSHLWGYAPFFSLANE 60
QY 61 SAISPDVPKGCRTVFQVLSRHRGARYPTSSASKAYSALIEAQKNATAFKKYAFLLKTYN 120
Db 61 SAISPDVPKGCRTVFQVLSRHRGARYPTSSASKAYSALIEAQKNATAFKKYAFLLKTYN 120
QY 121 YTLGADDLTPFGQGMVNSGKIFRYRKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGQGLVNSGKIFRYQYRESLTNRNIPFIASGSSRVIASGERFIEGFQST 180
QY 181 KLADPCANPHQASPVNIIPGAGYNWTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPCANPHQASPVNIIPGAGYNWTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCMCPDVTARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
Db 241 ARLEAHLPGVNLTDDEVNLMDCMCPDVTARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
QY 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
Db 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
QY 361 HNTMVSIFFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAVEMMOCEAEKEP 420
Db 361 HNTMVSIFFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAVEMMOCEAEKEP 420
QY 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECFA 467
Db 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECFA 467

```

RESULT 5

```

US-08-374-652C-2
; Sequence 2, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; FILING DATE: 29-APR-1987
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON

```

Sat Oct 27 15:25:53 2001

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07058
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
PS-08-374-652C-2

Query Match . 75.0%; Score 1852; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 7.9e-183;
Matches 347. Conservative 42; Mismatches 78; Indels

Qy	1	MGVFPVLLSTATLFGSTGTALCPGNSHUSCDTVDGGYOCFPEISHWCTYSPFSLADE	60
Db	1	MGVSAVLLPYLLAGVYTSGLAVFASRNQSTCDVDGQYOCFSETSHWGOYAPFFSLANE	60
Qy	61	SATSPDVPKCRVTEYQVLSRHCARGPYTSSAKYKAYSALTEAQKNATAFGKYAFKLTYN	120
Db	61	SATSPDVPACRVTFQAQLSRHCARGPYTESKGGKYSALIEEQONTVTTEDGKYAFKLTYN	120
Qy	121	YTLGADDLTFEGQQVNSGIKFYRYKALARKIVFFIRASGSDRVIASEAKFIEGFQSA	180
Db	121	YSLGADDLTFPEGEQVNSGIKFYQRYESLTRNIIFIRASSGSRVIASEKFIIEGFQST	180
Qy	181	KLADPGANHQASPVINVIPEGAGYNNTLDHGLCTAFTESELDGDDVEANFATVAFAPPR	240
Db	181	KLKDPRAQPGOSSPKIDVFISEASNNNTLDPGCTCFEEDSELADIVEANFATFAPSIR	240
Qy	241	ARLEAHPGVNLTDEDVNNLMDMCPDFTVARTSDATQLSPFCDLTFHDEWIQYDYLOSIG	300
Db	241	QRLNDLSGVTLTDETVYLLMDMCSPTDITSTSTVDTKLSPFCDLTFHDEWIHYDYLQSLK	300
Qy	301	KYYGYGAGNPLGPAGQGVGFVNELIARLTHSPVDHSTNTHTLDSNPATPPLNATLTADPS	360
Db	301	KYYGHGAGNPLGPQTQGVGYANELIARLTHSPFVDDTSSNHTLDSNPATPPLNSTLTADFS	360
Qy	361	HDNTMVSIFFGALGYNCTKPLSTTSVESIETDGYASASWTVPFAAAYVYEMMOCEAKEP	420
Db	361	HDNGIISILFGALGUYNGTKPLSTTTVENIQTDTGFSASWTVPFASRLYVEMMOCAQFEP	420
Qy	421	LVRVLVNDRVVPLHGCGVDKLGRCRDRDFEGLSFARSGGNWEECEFA	467
Db	421	LVRVLVNDRVVPLHGCPTDALGRCTRDSFVRGLSFARSGGDWAECSA	467

```

RESULT      6
US-08-151-574-32
; Sequence 32, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Seltin
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-151-574-32

```

Query Match	Score	DB 1: Length
74.98:	1848:	467:

Best Local Similarity	74.38;	pred: No. 2e-182;	
Matches	247;	Conservative	43; Mismatches
			78; Indels
			0; Gaps
			0;

Qy	1	MGVFEVLLSTATLFGSTSGTALPGRNHSCDTPVDGGYQCFPEISHLWGTYPGFFSLADE	50
Db	1	MGVSAVLLPLYLLSGVTSGI LVA PASRNQSCDTPVDGGYQCFSETSHLWGTYPGFFSLANE	60
Qy	61	SAISPDYPKGRVTVEQVLSRHCARYPTSSAKAYSALTEAQKNATAFKGYAFLEKTYN	120
Db	61	SVISPEVPAGCRVTFAQVLSRHGARTPTDSKGKYSALIEETQONATTEDGKYAFLEKTYN	120
Qy	121	YTLICADDLTPFGEOQVNSNGIKRYRYKALARKIVPFI RASGSDRVIASAEKPIEGFQSA	180
Db	121	YSLGADDLTPFGEOQLVNSGIKEGYQRYESLTRINIVPFI RASGSSRVIASAGKKPIEGFQST	180
Qy	181	KLADPGANPHQASPVINVIITPEGAGYNNLTDLGLCTAFTESESLGDDVEANFATVAFAPPIR	240
Db	181	KLKDPRAQPGQSSPKIDVILISEASSSNTLDPGCTCFVDESELADVEANFATVFPVSIR	240
Qy	241	ARLEAHLPGVNNLTDEDVYNNLMDMCPFDFTVARTSDATQLSPFCOLFTHDEWIQYDYLSQLS	300
Db	241	ORLENDLSGVTLTDEVTYLYMDMCSFTDITSSVDYTKLSPFCDLFTHDEWIQYDYLSQLSK	300

QY 301 KYGYGAGNPLGPAQGVGFVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYGHGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMVSIFPAGLVNCTKPLSTTSVESIETDGYASMTVPFAARAYVEMMOCEAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSAWTPPFASRLYVEMMOCEAEKEP 420
 QY 421 LVRVLVNDRVVPLHGGGVDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLVNDRVVPLHGGCPVDALGRCTRDSFVRGLSFARSGGDWAECEFA 467
 RESULT 7
 US-08-146-424-20
 ; Sequence 20, Application US/08146424
 ; Patent No. 5593963
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN COIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/146,424
 ; FILING DATE: 02-NOV-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KENNEDY, BILL
 ; REGISTRATION NUMBER: 33,407
 ; REFERENCE/DOCKET NUMBER: 44615-20011.24
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-146-424-20
 Query Match 74.9%; Score 1848; DB 1; Length 467;
 Best Local Similarity 74.3%; Pred. No. 2e-182;
 Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;
 QY 1 MGFWVLLSIATLFGSTSGTALGPRGNHSCDVTDDGGYQCFPEISHLWGTYSPPFFSLADE 60
 Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSSCDTVDGGYQCFSETHLWQYAPFESLANE 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKAYSALIEIAIQKNATAFKGYAPFLKTYN 120
 Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTDSKGYKYSALIEIEIQQNATTFDCKYAPFLKTYN 120
 QY 121 YTLGADDLTPFGQOMVNSGIKEFYRRYKALARKIVPFIRASGSDRVIASAEKFTGQSA 180

Db 121 YSLGADDLTPFGQELVNSGIKEFYRYSLTRNIVPFISSGSSRVIASCKKFTGQST 180
 QY 181 KLADPPGANPHQASVINVIIPGAGYNNITLDHGLCTAFESSELGDDVEANFTAYFAPPIR 240
 Db 181 KLADPPAQPGQSSPKIDVIVISEASSNNTLDPGTCTVFDESELADTVEANFTATFVPSIR 240
 QY 241 ARLEAHLPGVNLTDDEDVYNLMDMCPDFTVARTSDATQLSPCDLTFTHDEWIQDYDLSLG 300
 Db 241 QRLNDLSGVTLTDTEVTYLMDCSFDITSTIVDTKLSPCDLTFHDEWINYDYLQSLK 300
 QY 301 KYCYGAGNPLGPAQGVGFVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYGHGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMVSIFPAGLVNCTKPLSTTSVESIETDGYASMTVPFAARAYVEMMOCEAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSAWTPPFASRLYVEMMOCEAEKEP 420
 QY 421 LVRVLVNDRVVPLHGGGVDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLVNDRVVPLHGGCPVDALGRCTRDSFVRGLSFARSGGDWAECEFA 467
 RESULT 8
 US-08-693-709-2
 ; Sequence 2, Application US/08693709
 ; Patent No. 5770413
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN COIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/693,709
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146,424
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20011.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-813-5600
 ; TELEFAX: 415-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal

```
;
;
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1-23
; OTHER INFORMATION:
;
; US-08-693-709-2
;
; Query Match 74.9%; Score 1848; DB 1; Length 467;
; Best Local Similarity 74.3%; Pred. No. 2e-182;
; Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;
;
; QY 1 MGVEVLLSIAIFLSTGTSCTALPGRNHSCDVTGQYQCFPEISHLWCTYSPFFSLADE 60
; DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFSETSHLWQYAPFFSLANE 60
;
; QY 61 SATSPDPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGKYAFELKTYN 120
; DB 61 SVISPEVPAGCRVTFQVLSRHGARYPTDSKGGKYSALIEEQNATTFDGKYAFELKTYN 120
;
; QY 121 YTLGADDLTPFGQOQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
; DB 121 YSLGADDLTPFGQOQMVNSGKIFRYQYESLIRNIVPFIASGSDRVIASGKFIQFQST 180
;
; QY 181 KLADPGANPHQASPVNIVIPGAGYNNLTHGLCTAPESELGDDVEANFTAVFAPPIR 240
; DB 181 KLADPGANPHQASPVNIVIPGAGYNNLTHGLCTAPESELGDDVEANFTAVFAPPIR 240
;
; QY 241 ARLEAHLPGVNLTDVNVNMDMCPDFTVARTSDATQLSPPFCDLTHDEWIDYQLQSLG 300
; DB 241 ORLENDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPFCDLTHDEWIDYQLQSLK 300
;
; QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
; DB 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
;
; QY 361 HDNTWVSIFPAGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAEKEP 420
; DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVFPASRLYVEMMQCAEQEP 420
;
; QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
; DB 421 LVRVLVNDRVVPLHGCPVDALGRCTRDRSFVRGLSFARSGGNWEECFA 467
;
; RESULT 9
; US-08-419-448-32
; Sequence 32, Application US/08419448
; Patent No. 5863533
;
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gortcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
```

```
;
;
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-419-448-32
;
; Query Match 74.9%; Score 1848; DB 2; Length 467;
; Best Local Similarity 74.3%; Pred. No. 2e-182;
; Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;
;
; QY 1 MGVEVLLSIAIFLSTGTSCTALPGRNHSCDVTGQYQCFPEISHLWCTYSPFFSLADE 60
; DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFSETSHLWQYAPFFSLANE 60
;
; QY 61 SATSPDPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGKYAFELKTYN 120
; DB 61 SVISPEVPAGCRVTFQVLSRHGARYPTDSKGGKYSALIEEQNATTFDGKYAFELKTYN 120
;
; QY 121 YTLGADDLTPFGQOQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
; DB 121 YSLGADDLTPFGQOQMVNSGKIFRYQYESLIRNIVPFIASGSDRVIASGKFIQFQST 180
;
; QY 181 KLADPGANPHQASPVNIVIPGAGYNNLTHGLCTAPESELGDDVEANFTAVFAPPIR 240
; DB 181 KLADPGANPHQASPVNIVIPGAGYNNLTHGLCTAPESELGDDVEANFTAVFAPPIR 240
;
; QY 241 ARLEAHLPGVNLTDVNVNMDMCPDFTVARTSDATQLSPPFCDLTHDEWIDYQLQSLG 300
; DB 241 ORLENDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPFCDLTHDEWIDYQLQSLK 300
;
; QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
; DB 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF 360
;
; QY 361 HDNTWVSIFPAGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAEKEP 420
; DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVFPASRLYVEMMQCAEQEP 420
;
; QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
; DB 421 LVRVLVNDRVVPLHGCPVDALGRCTRDRSFVRGLSFARSGGNWEECFA 467
;
; RESULT 10
; US-08-819-825-3
; Sequence 3, Application US/08819825
; Patent No. 5866118
;
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58661180 No. 5866118th disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
;
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-819-825-3

```

```

Query Match 74.9%; Score 1848; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYCFEISHLWGTYSPFFSLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYCFEISHLWGTYSPFFSLADE 60

QY 61 SAISDPVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKLTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKLTYN 120

QY 121 YTLGADDLTPFGEQELVNSGKIFRYKALARKIYVPIFRASGSDRVIAAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEQELVNSGKIFRYKALARKIYVPIFRASGSDRVIAAEKFIQFQSA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVVISASNSNTLDPGCTCTVEDESLADTVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSPFCDLTFHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVYLLMDMCSFDISTSTVDTKLSPECDLTFHDEWIQDYLSLG 300

QY 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMVSIFGALYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVENMQCAEKEP 420
Db 361 HDNGIISILFALYNGTKPLSTTIVENITQDGFSSAWTVPFAARAYVENMQCAEKEP 420

QY 421 LVRVLVNDVVPLHGGCVKDLGRCKRDDFVGLSFARSGGNWEECA 467
Db 421 LVRVLVNDVVPLHGGCPVDALGRCTRDSFVRGLSFARSGGNWEECA 467

```

```

RESULT 11
US-09-163-642-3
; Sequence 3, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berk, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: No. 6221644o No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-163-642-3

```

```

Query Match 74.9%; Score 1848; DB 4; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYCFEISHLWGTYSPFFSLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYCFEISHLWGTYSPFFSLADE 60

QY 61 SAISDPVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKLTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKLTYN 120

QY 121 YTLGADDLTPFGEQELVNSGKIFRYKALARKIYVPIFRASGSDRVIAAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEQELVNSGKIFRYKALARKIYVPIFRASGSDRVIAAEKFIQFQSA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVVISASNSNTLDPGCTCTVEDESLADTVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSPFCDLTFHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVYLLMDMCSFDISTSTVDTKLSPECDLTFHDEWIQDYLSLG 300

QY 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMVSIFGALYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVENMQCAEKEP 420
Db 361 HDNGIISILFALYNGTKPLSTTIVENITQDGFSSAWTVPFAARAYVENMQCAEKEP 420

QY 421 LVRVLVNDVVPLHGGCVKDLGRCKRDDFVGLSFARSGGNWEECA 467
Db 421 LVRVLVNDVVPLHGGCPVDALGRCTRDSFVRGLSFARSGGNWEECA 467

```

```

RESULT 12
US-09-155-855-3

```

; Sequence 3, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Synichi
; APPLICANT: KANEKO, Synichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Aspergillus niger
; US-09-155-855-3

Query Match 74.6%; Score 1841; DB 4; Length 467;

Best Local Similarity 73.2%; Pred. No. 1.1e-181;
Matches 342; Conservative 49; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGVEVLLSLATLFGSTGALPGRNHSCDVTGQYCFPEISHLWGTYSPPFSLADE 60
DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYCFPEISHLWGTYSPPFSLANK 60
QY 61 SAISPDVPGKCRVTFVOVLSRHGARYPTSSASKAYASALIEIAOKNATATKGYAFKLTYN 120
DB 61 SAISPDVPGKCRVTFVOVLSRHGARYPTSSASKAYASALIEIAOKNATATKGYAFKLTYN 120
QY 121 YTLGADDLTPFGQGVNMGKIKYRRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGQGVNMGKIKYRRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFESSELGDDVEANFTAFVAPPTR 240
DB 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFESSELGDDVEANFTAFVAPPTR 240
QY 241 ARLEAHLPGVNLTDVNNLMDMCPDPTVARTSDATQSPFCDLFTHDEWIQDYQLSLG 300
DB 241 ARLEAHLPGVNLTDVNNLMDMCPDPTVARTSDATQSPFCDLFTHDEWIQDYQLSLG 300
QY 301 KYGAGNPLGPAQGVNFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYAFDS 360
DB 301 KYGAGNPLGPAQGVNFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYAFDS 360
QY 361 HDNTMVSIFFALGYNKTKPLSTSVSIBETDGYASWTVFAARAYVEMMOCEAEKEP 420
DB 361 HDNTMVSIFFALGYNKTKPLSTSVSIBETDGYASWTVFAARAYVEMMOCEAEKEP 420
QY 421 LVRVLVNDVVPVPLHGGCPDVALGRCTRODSFVKGSLFARSGGDNWEECEFA 467
DB 421 LVRVLVNDVVPVPLHGGCPDVALGRCTRODSFVKGSLFARSGGDNWEECEFA 467

RESULT 13

US-09-155-855-1
; Sequence 1, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Synichi
; APPLICANT: KANEKO, Synichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
; US-09-155-855-1

Query Match 72.9%; Score 1799; DB 4; Length 443;

Best Local Similarity 74.8%; Pred. No. 2.2e-177;
Matches 330; Conservative 47; Mismatches 64; Indels 0; Gaps 0;

QY 27 NSHSCDVTGQYCFPEISHLWGTYSPPFSLADESAISPDVPGKCRVTFVOVLSRHGARY 86
DB 3 NQSTCDVTGQYCFPEISHLWGTYSPPFSLANKSAISPDVPGKCRVTFVOVLSRHGARY 62
QY 87 PTSSASKAYSALIEIAOKNATATKGYAFKLTYNITLGADDLTPFGQGVNMGKIKFYRR 146
DB 63 PTSSASKAYSALIEIAOKNATATKGYAFKLTYNITLGADDLTPFGQGVNMGKIKFYRR 122
QY 147 YKALARKIVPFIASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPEGAGY 206
DB 123 YESLRTNIVPFIASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPEGAGY 182
QY 207 NNTLDHGLCTAFESSELGDDVEANFTAFVAPPTRARLEAHLPGVNLTDVNNLMDMCPF 266
DB 183 NNTLDHGLCTAFESSELGDDVEANFTAFVAPPTRARLEAHLPGVNLTDVNNLMDMCPF 242
QY 267 DTVARTSDATQSPFCDLFTHDEWIQDYQLSLGKIKYRRYKALARKIVPFIASGSDRV 326
DB 243 DTVARTSDATQSPFCDLFTHDEWIQDYQLSLGKIKYRRYKALARKIVPFIASGSDRV 302
QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFALGYNKTKPLSTTSV 386
DB 303 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFALGYNKTKPLSTTSV 362
QY 387 ESIEETDGYASWTVFAARAYVEMMOCEAEKEPILVRVLVNDVVPVPLHGGCPDVALGRCTR 446
DB 363 ENITQDGFSSAATVFPFASRMVEMMOCEAEKEPILVRVLVNDVVPVPLHGGCPDVALGRCTR 422
QY 447 DDFVGLSFAKSGGDNWEECEFA 467
DB 423 DDFVGLSFAKSGGDNWEECEFA 443

RESULT 14

US-09-155-855-2
; Sequence 2, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Synichi
; APPLICANT: KANEKO, Synichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 443

; TYPE: PRT

; ORGANISM: Aspergillus niger

; FEATURE:

; OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown
US-09-155-855-2

Query Match

Best Local Similarity 72.7%; Score 1794; DB 4; Length 443;

Matches 329; Conservative 47; Mismatches 62; Indels 0; Gaps 0;

QY 30 SCDVGGYQCFPEISHLWCTYSPFTSLADESALSDVPKRCRVTFVQVLSRHGARYPTS 89

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 TCDVGGYQCFSETSHLWGOYAPFSLANKSALSDVPAGCHVTFAQVLSRHGARYPTD 65

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 90 SASKAYSALIEATOKNATAPKGYAEFLKTYNYTLGADDLTPFGGEQMVNSGKIFRYRYKA 149

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 66 SKGKYKYSALIEEQOATTFEGKYAEFLKTYNYSLGADDLTPFGGEQELVNSGVKRYORYES 125

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 150 LARKIYVFFRAGSDRVIASAEKFIQFQSAKLADPGAPHQASPVINVIIPGAGYNNY 209

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 126 LTRNIVFFIRSSGSRVIAAGNKFIQFQSTKLKDPRAQFGQSSPKIDVYVISEASTSNNT 185

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 210 LHGLCTAFEESLGDVDEANTAVFAPPTIRARLEAHLPGVNLTDDEVDVNLMDMCPDVT 269

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 186 LDPGCTVFEDELADIEANTATVPISIRQLENDLSGVSLDTEVYLMDCSFDTI 245

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 270 ARTSDATQSLPDCDLTHEDWIOYDYLQSLGYGYGAGNPLGPAQGVFNELIARLTH 329

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 246 STVDTKLSPFCDLTHEWINYDYLQSLNKYGYGAGNPLGPTQGVYANELIARLTH 305

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 330 SPVODHTSNHLDSPATFELNATLYADFSDNTMVSIFALGLYNGTKPLSTTSVESI 389

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 306 SPVHDTSNHLDSPATFELNATLYADFSDNTMVSIFALGLYNGTKPLSTTSVAENI 365

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 390 EETDGYASWTVPFAARAYVENMQCEBAEKPEPLRVLVNDRVPLHCGGYDKLGRCKRDOF 449

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 366 TQTDGFSSAWTVPEASRMVEMMQCEQEPPLRVLVNDRVPLHCGPYDALGRCTRDSF 425

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 450 VEGLSFARSGGNWECFA. 467

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 426 VKGLSFARSGGDWECFA. 443

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

RESULT 15

US-08-819-825-2

; Sequence 2, Application US/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5;

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 4758,200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; US-08-819-825-2

Query Match

Best Local Similarity 56.0%; Score 1381; DB 2; Length 475;

Matches 281; Conservative 61; Mismatches 107; Indels 38; Gaps 11;

QY 1 MGVF-VVLLSIATLFGSTSGTALGP--RGNHSCDVTVDGGYQCFPEISHLWGYSPFSL 57

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 LGSFLVLLLOFSALL--TASPAIPFPWRKKHPND-----IARHWGQTSPPFSL 52

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 58 ADSAISDPVDPKRCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLK 117

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 53 AEVSEISPAVPKRCRVTFVQVLSRHGARYPTAHKSEVYAEELLQRIQDTATEFKGDAFLR 112

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 118 TYNITLGADDLTPFGGEQMVNSGKIFRYRYKALARKIVPIIRASGSDRVIASAEKFIQF 177

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 113 DAYHIGAONLTFEGEQMMESGRQFYHRYEQAREIVPFVRAAGSARVIASAEFFNRGF 172

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 178 QSAKLADPGAPHQASPVINVIIPGAGYNNLTDHGLCTAFEESLGDVDE--ANFTAVFA 236

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 173 QDAKDRDPSRKDOAEPLVINVIIEETGSNNLTDGLTCAAEAP--DPTQPAEFLQVFG 230

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 237 PPIRARLEAHLPGVNLTDDEVDVNLMDMCPDVTVARTSDAT----QLSPFCDLTHDEWIQ 292

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 231 PRVLKIKITKHPGVNLTDLEDVPLFMDLCPDVTG--SDPVLFPRLQSLSPFCHLTADDWA 288

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 293 YDYLQSLGYGYGAGNPLGPAQGVFNELIARLT-HSPVQDHTSTNHTLDSNPATFPL 351

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 289 YDYITLDKYYSHGGSAFGPSRGVGFVNELIARMTGNLFPVKDHTTVNHTLDDNPETFL 348

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 352 NATLYADFSDNTMVSIFALGLYNGTKPLSTTSVE--STEETDGYASWTVPFAARAYV 409

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 349 DAVLYADFSDNTMTGIFSAMGLYNGTKPLSTSKIQPTTGAAADGYAASWTVPFAARAYV 408

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 410 EMQOC-----EAEKEPLRVLVNDRVPLHCGGYDKLGRCKRDOFVEGLSARS 459

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 409 ELLRCETETSEEEEGEDEPFVRLVNDVPLHRCVDRWRCRRDEWIKGLTFAEQG 468

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 460 GNWECF 466

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 469 GHWDRCF 475

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

Search completed: October 26, 2001, 16:40:47

Job time: 4958 sec

